Blind Signature Waveform Estimation and Linear Multiuser Detection in Direct Sequence Code Division Multiple Access Systems

Vom Fachbereich 18 der
Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

von
M.Sc. **Keyvan Zarifi**
geboren in Khorramshahr, Iran

Referent:	Prof. Dr. Alex B. Gershman
Korreferent:	Prof. Dr. Nikos D. Sidiropoulos
Tag der Einreichung:	Jan. 24, 2007
Tag der müundlichen Prüfung:	May 15, 2007
Erscheinungsjahr:	2007

D17
Darmstädter Dissertationen

To Baharak
and to my mother

Acknowledgments

First and foremost I would like to thank my best friend and my beloved wife Baharak for her love, friendship, understanding, and unconditional support throughout the years that we have been together. This thesis is just a tiny part of what I have because of her.

I was very lucky that I got the chance to work under supervision of Prof. Alex Gershman. His comments, suggestions, and discussions were always inspiring and what I learned from his professional manner is undoubtedly a major achievement of my graduate studies. His vision and his trust have helped me to develop the self-confidence required for any Ph.D. student who endeavors to become an independent researcher.

Many thanks to all of my colleagues both in Darmstadt University of Technology and my former university of Dusiburg-Essen for building lovely environments to work in. I would like to express my special gratitude to Prof. Thomas Kaiser, Prof. Andreas Czylwik, and our friendly and very helpful secretaries Mrs. Petra Kaiser in Dusiburg-Essen University and Mrs. Marlis Gorecki in Darmstadt.

I wish to give my special thanks to my long-term colleague Prof. Sergiy Vorobyov for his discussions and my friend Dr. Mohammad Amin Zia for his support and generous help.

I could not be where I am today without my brother Siamak and my lovely sisters Franak and Zoya. Last, but certainly far from the least, I send my purest love to my mother whom the words are humbled to thank.

Darmstadt, Germany — Keyvan Zarifi
Jan. 2007

Zusammenfassung

Direct-sequence code-division multiple-access (DS-CDMA) ist eine weit verbreitete Übertragungstechnik, die in drahtlosen Kommunikationssystemen der dritten Generation sowie darüber hinaus verwendet wird. Lineare Mehrbenutzerdetektionsverfahren mit geringem Rechenaufwand können die Kapazität von DS-CDMA Systemen wesentlich verbessern, indem nicht nur die Signatur des users-of-interest, sondern auch die der interferierenden Nutzer berücksichtigt wird. Da jedoch spreizspektrum DS-CDMA Signale meist durch frequenzselektives Fading gestört werden, stellt die Fehlanpassung zwischen angenommener und tatsächlicher Signatur des Benutzers empfängerseitig eine typische Herausforderung dar. Da Fehlanpassungen der Signatur die Performance des Mehrbenutzerempfängers erheblich verschlechtern können, werden vor der Mehrbenutzerdetektion gute Verfahren zur Signaturschätzung benötigt. Unter den verschiedenen Signaturschätzverfahren sind insbesondere die bandbreiteneffizienten blinden Algorithmen von besonderem Interesse, da sie keine Übertragung von Trainingssymbolen benötigen. Diese Tatsache führte zum Thema dieser Doktorarbeit, in der fortschrittliche blinde Signaturschätzverfahren und Mehrbenutzerdetektionstechniken für DS-CDMA Systeme entwickelt und analysiert werden.

Der erste Teil dieser Arbeit beschäftigt sich mit der blinden Schätzung der Signatur. Wir entwickeln zuerst ein neuartiges Verfahren zur unterraumbasierten Schätzung der Signatur für binary phase shift keying (BPSK) Signale bei weiem Rauschen. Anders als die bekannten konventionellen Signaturschätzverfahren verwendet unser Algorithmus das empfangene Signal zusammen mit seinem komplex konjugierten, um die Performance der Schätzung wesentlich zu verbessern. Wir analysieren dann die Performance von zwei bekannten unterraumbasierten Schätzverfahren für den Fall von unbekanntem korrelierten Rauschen. Am Ende stellen wir dann eine neue Technik zur Signaturschätzung bei korreliertem Rauschen vor, die anders als die beiden vorherigen Schätzer bei beliebigen Signalkonstellationen verwendet und deutlich einfacher implementiert werden kann.

Der zweite Teil dieser Doktorarbeit beschäftigt sich mit blinden linearen Mehrbenutzerempfängern. In diesem Abschnitt entwickeln wir zuerst einen robusten blinden minimum mean-square error (MMSE) Empfänger, der explizit sowohl Fehler bei der Schätzung der Signatur als auch die Kovarianzmatrix des Signals berücksichtigt. Als Nächstes analysieren wir die signal-to-interference-plus-noise ratio (SINR) Performance des weit verbreiteten blinden minimum output energy (MOE) Empfängers in groen zufälligen DS-CDMA Kommunikationssystemen. Auerdem werden die asymptotischen Eigenschaften des Capon Empfängers

und des blinden Capon Kanalschätzers sowohl für das Uplink als auch für das Downlink Szenario im Detail untersucht.

Unsere Untersuchung bekannter blinder Signaturschätzverfahren und Mehrbenutzerdetektionsalgorithmen zeigen einige wichtige Eigenschaften dieser Verfahren, die in der Literatur bisher fehlen. Auerdem genügen unsere neuen Algorithmen den Anforderungen der sich weiterentwickelnden Kommunikationssystemstandards für verlässliche und dennoch einfache Detektionsstrategien. Daher erwarten wir, dass unsere Ergebnisse die Anwendung blinder Mehrbenutzerdetektion in zukünftigen Generationen von DS-CDMA Kommunikationssystemen fördern werden.

Abstract

Direct-sequence code-division multiple-access (DS-CDMA) technology is a popular communication scheme adopted for the third generation (3G) wireless communication systems and beyond. Computationally efficient linear multiuser detection techniques can significantly improve the capacity of DS-CDMA systems by taking into account not only the signature of the user-of-interest but also those of the interfering users. However, as the spread spectrum DS-CDMA signals are usually subject to frequency-selective fading, a mismatch between the presumed and the actual user signatures is a typical challenge at the receiver side. As such signature mismatch may have a substantially detrimental effect on the performance of multiuser receivers, it is required to use an accurate signature estimation technique prior to the subsequent multiuser detection procedure. Among various signature estimation approaches, the bandwidth efficient *blind* algorithms that do not require transmission of any training symbols are of significant interest. This fact have motivated the topic of this thesis where advanced blind signature estimation and multiuser detection techniques for DS-CDMA systems are developed and studied.

The first part of the thesis is devoted to blind signature estimation. We first develop a novel subspace-based signature estimation technique for binary phase shift keying (BPSK) transmitted signals in a white noise scenario. Unlike the earlier conventional signature estimation techniques, our method exploits the received signal jointly with its complex conjugate to substantially improve the estimation performance. We then analyze the performance of two well-known subspace-based estimation techniques developed for the case of unknown correlated noise. Finally, we propose a new signature estimation technique for this case that, unlike the latter two estimation algorithms, can be applied to an arbitrary signal constellation and enjoys much simpler implementation.

The second part of the thesis is devoted to blind linear multiuser receivers. In this part of our work, we first develop a robust blind minimum mean-square error (MMSE) receiver that explicitly accounts for both the errors in the estimated signature and the sample data covariance matrix. Next, we analyze the signal-to-interference-plus-noise ratio (SINR) performance of the popular blind minimum output energy (MOE) receiver in large random DS-CDMA communication systems. The asymptotic properties of the Capon receiver and the blind Capon channel estimate are also studied in detail for both uplink and downlink scenarios.

Our study of existing blind signature estimation and multiuser detection algorithms

reveal some important properties of these techniques that have been missing in the literature. Moreover, our novel algorithms meet the requirements of the evolving communication network standards for highly reliable yet simple detection strategies. As such, we expect that our results will facilitate the application of blind multiuser detection to the future generations of DS-CDMA communication systems.

Table of Contents

Acknowledgments **iii**

Zusammenfassung **v**

Abstract **vii**

List of Figures **xi**

Notation **xvii**

Acronyms **xix**

1 Introduction **1**
- 1.1 Background . . . 1
- 1.2 Overview and the thesis contributions . . . 2
- 1.3 Assumptions and signal model . . . 6
  - 1.3.1 Assumptions . . . 6
  - 1.3.2 Signal model . . . 7

2 Signature estimation in the presence of white noise **11**
- 2.1 Introduction . . . 11
- 2.2 ISI-free signal model . . . 13
- 2.3 The LX algorithm . . . 14
  - 2.3.1 Temporally oversampled version of the LX algorithm . . . 17
- 2.4 The proposed algorithm for BPSK-modulated systems . . . 19
  - 2.4.1 Identifiability condition . . . 23
  - 2.4.2 Performance analysis . . . 25
  - 2.4.3 Extension to temporally oversampled received signals . . . 32
- 2.5 Simulations . . . 33
- 2.6 Chapter summary . . . 40

3 Signature estimation in the presence of correlated noise 41
3.1 Introduction . . . 41
3.2 Blind channel estimation . . . 43
3.2.1 WP algorithm . . . 43
3.2.2 BP algorithm . . . 45
3.3 Performance analysis . . . 46
3.3.1 Performance analysis of the WP algorithm . . . 46
3.3.2 Performance analysis of the BP algorithm . . . 54
3.4 Simulations . . . 60
3.5 Chapter summary . . . 66

4 Signature estimation in the presence of wide-sense stationary noise 69
4.1 Introduction . . . 69
4.2 The proposed technique . . . 71
4.3 Performance analysis . . . 75
4.4 Temporally oversampled version of the proposed algorithm . . . 83
4.5 Simulations . . . 85
4.6 Chapter summary . . . 90

5 Robust blind MMSE multiuser detection 93
5.1 Introduction . . . 93
5.2 Background . . . 95
5.2.1 Linear receivers . . . 95
5.2.2 MMSE receiver . . . 96
5.3 Robust multiuser detection . . . 98
5.3.1 Formulation . . . 98
5.3.2 Summary of the proposed multiuser receiver . . . 106
5.3.3 Relationship to the SOCP multiuser receiver . . . 106
5.4 Simulations . . . 108
5.5 Chapter summary . . . 115

6 Asymptotic performance analysis of MOE receivers in large DS-CDMA systems 117
6.1 Introduction . . . 117
6.2 Background . . . 120
6.3 Asymptotic analysis . . . 124
6.4 Channel distribution effect on the receiver performance . . . 137
6.5 Simulations . . . 141
6.6 Chapter summary . . . 147

7 High SNR extension to asymptotic performance analysis of MOE receivers **149**
7.1 Performance analysis 149
7.2 Simulations 164
7.3 Chapter summary 168

8 Conclusions and future work **177**
8.1 Conclusions 177
8.2 Future work 181

Bibliography **183**

Résumé **197**

List of Figures

2.1 Channel estimation MSEs versus the number of users. 34
2.2 Channel estimation MSEs versus the channel order. 34
2.3 Channel estimation MSEs versus the SNR. 35
2.4 Channel estimation MSEs versus the observation length. 36
2.5 Signature estimation MSEs versus the SNR for $\check{L} = 5, 6, 7, 10$. 37
2.6 BERs of the MMSE receivers versus the number of users. 38
2.7 Channel estimation MSEs versus the number of users in the the presence of oversampling. 39
2.8 Channel estimation MSEs versus the channel order in the presence of oversampling. 40

3.1 The MSE of the estimated channel versus SNR. The WP algorithm. 61
3.2 The MSE of the estimated channel versus number of data samples. The WP algorithm. 61
3.3 The MSE of the estimated channel versus SNR at the first antenna for different values of SNR at the second antenna. The WP algorithm. 62
3.4 The MSE of the estimated channel versus SNR. The BP algorithm. 63
3.5 The MSE of the estimated channel versus number of data samples. The BP algorithm. 64
3.6 The MSE of the estimated channel versus SNR for $\tilde{\mathbf{\Gamma}}_v = \text{diag}\{20, 5, 3\}$ and different matrices $\tilde{\mathbf{U}}_v$. The BP algorithm. 65
3.7 MSEs of the estimated channel versus SNR for $e_o = 28$ and different matrices $\mathbf{\Sigma}_v$. The BP algorithm. 66

3.8 MSEs of the estimated channel versus SNR in the white noise environment. The LX, WP, and BP algorithms. 67
3.9 MSEs of the estimated channel versus L for $L_c = 40$ and $L_c = 80$. The BP algorithm. 67

4.1 MSEs of the estimated channel versus the number of data samples N for the first noise model. 85
4.2 MSEs of the estimated channel versus SNR for the first noise model. 86
4.3 MSEs of the estimated channel versus number of users for the first interference model. 87
4.4 MSEs of the estimated channel versus number of data samples N for the second interference model. 88
4.5 MSEs of the estimated channel versus SNR for the second interference model. 89
4.6 MSEs of the estimated channel versus SNR for the third interference model 90

5.1 BERs versus SNR. First example. 108
5.2 BERs versus N. First example. 109
5.3 BERs of the proposed receiver versus $\varepsilon/\|\mathbf{c}_1\|$. First example. 110
5.4 BERs versus δ. First example. 111
5.5 BERs versus SNR. Second example. 112
5.6 BERs versus N. Second example. 113
5.7 The BER of the proposed receiver versus $\varepsilon/\sqrt{L_c}$. Second example. 114
5.8 BERs versus SNR. Third example. 115

6.1 $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{1}))$ versus α in the downlink case. 142
6.2 $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{1}))$ versus α in the uplink case. 143
6.3 $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ versus α in the downlink case. 143
6.4 $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ versus α in the uplink case. 144
6.5 Δ versus α in the downlink (upper subplots) and uplink (lower subplots) cases. 145
6.6 MSE between $\mathbf{h}_{\mathrm{C}}$ and the normalized version of $\mathbf{h}_1$ versus $\sigma_{\mathbf{h}}^2/\sigma^2$. 146
6.7 $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus $\sigma_{\mathbf{h}}^2/\sigma^2$. 147
6.8 Δ versus the actual channel length $\tilde{L}$ when $L = 15$ is assumed in the Capon receiver. 148

7.1 MSE of the Capon channel estimate versus $\sigma_{\mathbf{h}}^2/\sigma^2$ for different values of α in the uplink case. 166
7.2 SINR($\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})$) versus $\sigma_{\mathbf{h}}^2/\sigma^2$ for different values of α in the uplink case. . . . 168
7.3 SINR($\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1)$) versus $\alpha > 1$ in the uplink case for $\sigma_{\mathbf{h}}^2/\sigma^2 = 10$ dB (upper subplots), $\sigma_{\mathbf{h}}^2/\sigma^2 = 30$ dB (middle subplots), and $\sigma_{\mathbf{h}}^2/\sigma^2 = +\infty$ (lower subplots) in the uplink case. 170
7.4 SINR($\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})$) versus $\alpha > 1$ in the uplink case for $\sigma_{\mathbf{h}}^2/\sigma^2 = 10$ dB (upper subplots), $\sigma_{\mathbf{h}}^2/\sigma^2 = 30$ dB (middle subplots), and $\sigma_{\mathbf{h}}^2/\sigma^2 = +\infty$ (lower subplots) in the uplink case. 171
7.5 C($\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1)$) versus α in the uplink case for $\sigma_{\mathbf{h}}^2/\sigma^2 = 10$ dB (upper subplots), $\sigma_{\mathbf{h}}^2/\sigma^2 = 20$ dB (middle subplots), and $\sigma_{\mathbf{h}}^2/\sigma^2 = 40$ dB (lower subplots) in the uplink case. 172
7.6 C($\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})$) versus α in the uplink case for $\sigma_{\mathbf{h}}^2/\sigma^2 = 10$ dB (upper subplots), $\sigma_{\mathbf{h}}^2/\sigma^2 = 20$ dB (middle subplots), and $\sigma_{\mathbf{h}}^2/\sigma^2 = 40$ dB (lower subplots). . . 173
7.7 C($\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1)$) versus α in the downlink case for $\sigma_{\mathbf{h}}^2/\sigma^2 = 10$ dB (upper subplots), $\sigma_{\mathbf{h}}^2/\sigma^2 = 20$ dB (middle subplots), and $\sigma_{\mathbf{h}}^2/\sigma^2 = 40$ dB (lower subplots). . . 174
7.8 C($\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})$) versus α in the downlink case for $\sigma_{\mathbf{h}}^2/\sigma^2 = 10$ dB (upper subplots), $\sigma_{\mathbf{h}}^2/\sigma^2 = 20$ dB (middle subplots), and $\sigma_{\mathbf{h}}^2/\sigma^2 = 40$ dB (lower subplots). . . 175

Notation

$\mathrm{E}\{\cdot\}$	statistical expectation operation
$\texttt{dim}\{\cdot\}$	dimension of a subspace
$\texttt{range}(\cdot)$	range of a matrix
$\texttt{rank}(\cdot)$	rank of a matrix
$\texttt{null}(\cdot)$	null space a matrix
$\lambda_{\min}(\cdot)$	smallest eigenvalue of a matrix
$\lambda_{\max}(\cdot)$	largest eigenvalue of a matrix
$\mho(\cdot)$	normalized eigenvector associated with the smallest eigenvalue
$\Omega(\cdot)$	normalized eigenvector associated with the largest eigenvalue
$\|\cdot\|_F$	Frobenius norm of a matrix
$\|\cdot\|$	second norm of a vector or a matrix
$\|\cdot\|$	absolute value
δ_{ij}	Kronecker delta
$\mathbf{I}$	identity matrix
$\mathbf{I}_n$	$n \times n$ identity matrix
$\mathbf{J}$	matrix with ones on the main anti-diagonal and zeros elsewhere
$\mathrm{tr}(\cdot)$	trace of a matrix
$[\mathbf{v}]_i$	ith entry of vector $\mathbf{v}$
$[\mathbf{A}]_{i\bullet}$	ith row of matrix $\mathbf{A}$

$[\mathbf{A}]_{\bullet j}$	jth column of matrix $\mathbf{A}$
$[\mathbf{A}]_{ij}$	(i,j)th entry of matrix $\mathbf{A}$
$\mathbf{0}$	zero matrix
$\text{diag}(\mathbf{v})$	diagonal matrix whose diagonals are the entries of $\mathbf{v}$
$\mathbf{1}$	a vector with all entries equal to one
$\mathbf{e}_j$	a vector whose jth entry is one and the rest are zero
$\mathbf{A}^\dagger$	pseudo-inverse of matrix $\mathbf{A}$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker matrix product
$(\cdot)^T$	transpose
$(\cdot)^*$	complex conjugate
$(\cdot)^H$	transpose complex conjugate (Hermitian) operation
$\approx$	approximately equal
$\doteq$	equal up to the first order
$\text{Re}(\cdot)$	real part
$\text{Im}(\cdot)$	imaginary part
$\in$	membership
$\mathcal{S}_1 \cap \mathcal{S}_2$	intersection between subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$
$\text{sgn}[\cdot]$	sign function
$\text{I}\{\cdot\}$	indicator function
$\xrightarrow{a.s.}$	almost sure convergence
$\xrightarrow{e.a.s.}$	element-wise almost sure convergence
$\xrightarrow{i.p.}$	convergence in probability
$\xrightarrow{e.i.p.}$	element-wise convergence in probability

Acronyms

AR	auto-regressive
ASK	amplitude-shift keying
BER	bit error rate
BPSK	binary phase-shift keying
CCD	canonical correlation decomposition
CDMA	code division multiple-access
DOA	direction-of-arrival
DS-CDMA	direct-sequence code division multiple-access
FIR	finite impulse response
ICI	inter-chip-interference
INR	interference-to-noise ratio
ISI	inter-symbol-interference
LS	least-square
MIMO	multiple-input multiple-output
ML	maximum likelihood
MMSE	minimum mean-square error
MOE	minimum output energy
MSE	mean-square error
OQAM	offset quadrature-amplitude modulation

QPSK	quadriphase-shift keying
RF	radio frequency
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
SOCP	second-order cone programming
SVD	singular value decomposition
UMTS	universal mobile telecommunications system

Chapter 1

Introduction

1.1 Background

The ever growing demand for higher transmission rates to a larger number of users in 3G and beyond generations of wireless communication systems has strongly motivated the application of multiple-access communication systems [23], [31], [84]. DS-CDMA technology represents a very popular multiple-access scheme in which the user signals are convolved with distinct spreading codes while simultaneously share the same channel bandwidth [25], [37], [58], [64], [94]. It has been shown that the major factor which limits the performance of the DS-CDMA systems is the inherent co-channel interference [11], [31], [36], [91]. This observation has sparked an extensive amount of research on developing advanced and yet computationally affordable linear multiuser receivers that make use of not only the signal structure of the user-of-interest, but also that of the co-channel interference to significantly improve the detection performance [32], [84].

In spite of the elegant theory supporting the performance improvements promised by linear multiuser detection schemes, delivering these promises is a challenging task in the real world scenario where the spread-spectrum DS-CDMA signals are often subject to an unknown frequency-selective channel. Passing the DS-CDMA signals through such a channel results in an unknown distortion of the spreading codes, and, consequently, mismatch between the presumed and actual signature waveforms (that is, the spreading code convolved with the channel impulse response) at the receiver side. As a mismatched signature

waveform can result into a substantial degradation of the performance of multiuser receivers [31], [91], [101], [117], signature waveform estimation at the receiver side is an important prerequisite of the subsequent multiuser detection procedure. A classical approach to the signature waveform estimation problem is to transmit a sequence of training symbols which is known to the receiver, and then to estimate the signature waveform at the receiver side by computing the correlation between the training sequence and the received data [50], [54]. Since CDMA communication systems are typically employed in highly dynamic environments, a reliable signature waveform estimation requires frequent transmission of the training sequence. This, in turn, reduces the information transmission rate and wastes the available channel bandwidth [31], [91]. To overcome this drawback, considerable efforts have been devoted to develop the so-called blind signature estimation techniques that do not require any training symbols and are entirely based on the user spreading codes [6], [46], [47], [72], [76], [78], [89], [91], [101], [102], [111].

In this thesis, we develop and study advanced techniques in blind signature estimation and linear multiuser detection for DS-CDMA systems. In Chapters 2-4 we study the blind signature estimation problem, while in Chapters 5-7 we consider the blind linear multiuser detection techniques. In this thesis, the blind multiuser detection is referred to the class of multiuser detection techniques that do not rely on any training sequences to estimate the user signatures. The contributions of the thesis can be summarized as follows.

1.2 Overview and the thesis contributions

Chapter 2: In this chapter we first have a brief overview on the conventional white-noise assumption based signature estimation technique of [47]. Then, we develop a novel signature estimation algorithm for BPSK-modulated signals that makes use of both the received signal and its complex conjugate to effectively double the observation space, and, consequently, improve the estimation performance. We also prove the necessary and sufficient condition for channel identifiability and derive the MSE of the estimated channel. An extension of our technique to the temporal oversampled data is also presented. The results of this chapter

have been reported in the following publications:

- K. Zarifi and A. B. Gershman, "Blind subspace-based signature waveform estimation in BPSK-modulated DS-CDMA systems with circular noise," *IEEE Transactions on Signal Processing*, vol. 54, pp. 3592-3602, Sep. 2006.

- K. Zarifi and A. B. Gershman, "Enhanced blind subspace-based signature waveform estimation in CDMA systems with circular noise," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, USA, March 2005, pp. 937-940.

Chapter 3: In this chapter, we analyze the performances of two popular signature estimation techniques of [11] and [90] proposed for scenarios with unknown correlated noise. We show how the MSE performances of these techniques depend on the parameters such as the number of data samples, the spreading factor, the channel length, the power of the user-of-interest as well as the directions of the eigenvectors of the noise covariance matrix. We also prove that, if noise is white, then the MSE estimation performances of the algorithms proposed in [11], [47], and [90] are identical. The results of this chapter have been published in the following papers:

- K. Zarifi and A. B. Gershman, "Performance analysis of blind subspace-based signature estimation algorithms for DS-CDMA systems with unknown correlated noise," accepted for publication in *EURASIP Journal on Applied Signal Processing*, special issue on *Advances in Subspace-Based Techniques for Signal Processing and Communications*, to appear in 2007.

- K. Zarifi and A. B. Gershman, "Performance analysis of subspace-based signature waveform estimation algorithms for DS-CDMA systems with unknown correlated noise," in *Proc. IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, New York, USA, June 2005, pp. 600-604.

Chapter 4: This chapter proposes a novel subspace-based signature estimation technique for unknown wide-sense stationary correlated noise scenario. Unlike the techniques

of [11] and [90], our proposed technique does not require an auxiliary receive antenna and is also applicable to arbitrary signal constellations. We obtain the necessary and sufficient channel identifiability condition and derive the MSE estimation performance of our technique. We also show how the performance of our algorithm can be improved using the temporal oversampling of the received signal. The following papers report the results of this chapter:

- K. Zarifi and A. B. Gershman, "Blind subspace-based signature estimation in DS-CDMA systems with unknown wide-sense stationary interference," accepted for publication in *IEEE Transactions on Signal Processing*, to appear in July 2007.

- K. Zarifi and A. B. Gershman, "Subspace-based blind channel estimation in DS-CDMA systems with unknown wide-sense stationary interference," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006, pp. 121-124.

Chapter 5: In this chapter we develop a new blind robust MMSE receiver which accounts for errors of both the presumed signature waveform as well as the estimated data covariance matrix. A computationally efficient approach to derive our robust receiver is also presented. The results of this chapter have been published in the following papers:

- K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z.-Q. Luo, "Robust blind multiuser detection based on the worst-case performance optimization of the MMSE receiver," *IEEE Transactions on Signal Processing*, vol. 53, pp. 295-305, Jan. 2005.

- K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z.-Q. Luo, "Robust blind multiuser detection based on worst-case MMSE performance," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* Montreal, Canada, May 2004, pp. 897-900.

Chapter 6: In this chapter we analyze the asymptotic SINR performances of a general blind MOE receiver and its optimized version (known as the Capon receiver) in large random DS-CDMA communication systems. The asymptotic properties of the blind Capon channel

estimate are also discussed. We show that if either the user channels are frequency-flat or frequency-selective with zero-mean circular and independent channel coefficients, then the Capon channel estimation error converges to zero and the SINR performance of the Capon receiver approaches its maximum value. As having independent channel coefficients is commonplace in the uplink while is impossible in the downlink, we conclude that Capon estimation/detection technique is more useful for the uplink rather than for the downlink. The results of this chapter are presented in the following papers:

- K. Zarifi and A. B. Gershman, "Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems–Part I: General SNR framework," submitted to *IEEE Transactions on Signal Processing*.

- K. Zarifi and A. B. Gershman, "Large system performance analysis of minimum output energy receivers for DS-CDMA systems," submitted to *IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Helsinki, Finland, June 2007.

Chapter 7: This chapter continues the studies of Chapter 6 for the high SNR regime. It is shown that, if the spreading factor is larger than the number of users, then, as SNR goes to infinity, the Capon channel estimate converges to the channel of the user-of-interest and, moreover, the SINR of the MOE receiver goes to infinity. When the number of users is larger than the spreading factor, the SINR is shown to remain bounded. The asymptotic SINR expression of the MOE receiver is obtained for such a scenario. The efficiency and the asymptotic efficiency expressions of the MOE receiver are also derived and the corresponding approximations are presented for the case when the channel vectors are zero-mean circular and independent random vectors. The following papers are based on the results of this chapter:

- K. Zarifi and A. B. Gershman, "Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems–Part II: High SNR analysis," submitted to *IEEE Transactions on Signal Processing*.

- K. Zarifi and A. B. Gershman, "High SNR performance analysis of blind minimum output energy receivers in large DS-CDMA systems," accepted for *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Honolulu, USA, Apr. 2007.

Chapter 8: This chapter contains concluding remarks. We also propose some future work to further extend/improve the results of this thesis.

1.3 Assumptions and signal model

1.3.1 Assumptions

Throughout this thesis we study the synchronous CDMA systems. Note that the synchronous case is mainly considered for the sake of notational brevity. In fact, while extending our study in Chapter 6 and Chapter 7 to the asynchronous scenario requires some advanced results from random matrix theory, generalizing our results in Chapter 2 through Chapter 5 to the asynchronous systems is straightforward [89], [90]. Throughout the thesis, we also use the following three common assumptions:

A1 The chip sequence period is the same as the symbol period. This corresponds to the so-called short spreading code case [32].

A2 The user channels are quasi-static, i.e. the corresponding impulse channel responses are fixed during the observation period. Moreover, the lengths of the user channels are much less than the symbol period [14], [32], [47].

A3 Noise is a zero-mean random process. The transmitted symbols of each user are zero-mean unit-variance i.i.d. random variables, independent from that of the other users, and independent from noise [47].

We also assume in Chapter 2 and Chapters 5-7 that the noise is white. Other assumptions that apply only to some parts of the thesis will be introduced whenever required.

1.3.2 Signal model

The received continuous-time baseband signal of a K-user synchronous DS-CDMA system can be modelled as [78], [84]

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{k=1}^{K} A_k b_k(m) w_k(t - mT_s) + v(t) \tag{1.1}$$

where T_s is the symbol period, A_k, $b_k(m)$, and $w_k(t)$ denote the signal amplitude, the mth data symbol, and the signature waveform of the kth user, respectively, and $v(t)$ is the noise waveform.

Let L_c be the spreading factor and

$$\mathbf{c}_k = \big[c_k[0], c_k[1], \ldots, c_k[L_c - 1]\big]^T \tag{1.2}$$

denote the spreading sequence associated with the kth user. According to assumptions **A1** and **A2**, the signature waveform of this user is given by [47]

$$w_k(t) = \sum_{l=0}^{L_c-1} c_k[l] h_k\big(t - lT_c\big) \tag{1.3}$$

where $h_k(t)$ is the channel of the kth user that also includes the transmitter and receiver filters effect and

$$T_c = T_s/L_c \tag{1.4}$$

is the chip period. Let $[0, \alpha_k T_c]$ be the support of channel $h_k(t)$, where

$$L - 1 \leq \max\{\alpha_1, \ldots, \alpha_K\} < L \tag{1.5}$$

and L is a positive integer. Note that from **A2** we have

$$L \ll L_c. \tag{1.6}$$

Sampling (1.1) at the interval corresponding to the nth transmitted symbols with the period of T_c, the received sampled data vector can be written as [78]

$$\underline{\mathbf{x}}(n) = \sum_{k=1}^{K} A_k b_k(n) \underline{\mathbf{w}}_k + \sum_{k=1}^{K} A_k b_k(n-1) \underline{\tilde{\mathbf{w}}}_k + \underline{\mathbf{v}}(n) \tag{1.7}$$

where

$$\underline{\mathbf{x}}(n) = [x(nT_s), x(nT_s + T_c), \ldots, x(nT_s + (L_c - 1)T_c)]^T \tag{1.8}$$

$$\underline{\mathbf{w}}_k = [w_k(0), w_k(T_c), \ldots, w_k((L_c - 1)T_c)]^T \tag{1.9}$$

$$\underline{\tilde{\mathbf{w}}}_k = [w_k(L_cT_c), \ldots, w_k((L_c + L - 2)T_c), 0, \ldots, 0]^T \tag{1.10}$$

$$\underline{\mathbf{v}}(n) = [v(nT_s), v(nT_s + T_c), \ldots, v(nT_s + (L_c - 1)T_c)]^T. \tag{1.11}$$

Note that the second summation in (1.7) represents the ISI-contaminated part of the received signal. From (1.3), it follows that $\underline{\mathbf{w}}_k$ can be written in either of the following forms [14], [78]:

$$\underline{\mathbf{w}}_k = \underbrace{\begin{bmatrix} c_k[0] & 0 & \cdots & 0 \\ c_k[1] & c_k[0] & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ & & & c_k[0] \\ \vdots & \vdots & \ddots & \vdots \\ c_k[L_c-1] & c_k[L_c-2] & \cdots & c_k[L_c-L] \end{bmatrix}}_{\underline{\mathcal{C}}_k} \mathbf{h}_k = \underbrace{\begin{bmatrix} h_{k,0} & 0 & \cdots & 0 \\ \vdots & \ddots & & \\ h_{k,L-1} & \ddots & \ddots & \vdots \\ 0 & \ddots & h_{k,0} & \\ \vdots & \ddots & \ddots & \vdots \ \ddots \ 0 \\ 0 & \cdots & 0 & h_{k,L-1} \ \ldots \ h_{k,0} \end{bmatrix}}_{\mathcal{H}_k} \mathbf{c}_k \tag{1.12}$$

where

$$\mathbf{h}_k = [h_{k,0}, h_{k,1}, \ldots, h_{k,L-1}]^T \tag{1.13}$$

and

$$h_{k,i} = h_k(iT_c). \tag{1.14}$$

Using (1.3), it can also be shown that [14], [78]

$$\underline{\tilde{\mathbf{w}}}_k = \underbrace{\begin{bmatrix} 0 & c_k[L_c-1] & \dots & c_k[L_c-L+1] \\ \vdots & 0 & \ddots & \vdots \\ & \vdots & \ddots & c_k[L_c-1] \\ & & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}}_{\underline{\tilde{\mathbf{C}}}_k} \mathbf{h}_k = \underbrace{\begin{bmatrix} 0 \cdots & 0 & h_{k,L-1} & \dots & h_{k,1} \\ & \ddots & 0 & \ddots & \vdots \\ & & & \ddots & h_{k,L-1} \\ \vdots & \vdots & & \ddots & 0 \\ & & & & \vdots \\ 0 \cdots & 0 & \cdots & & 0 \end{bmatrix}}_{\tilde{\boldsymbol{\mathcal{H}}}_k} \mathbf{c}_k. \quad (1.15)$$

In linear multiuser detection schemes, we often require the knowledge of the data covariance matrix

$$\underline{\mathbf{R}} = \mathrm{E}\{\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H\}. \quad (1.16)$$

From (1.7), (1.12), (1.15) along with (2.11), it directly follows that

$$\begin{aligned} \underline{\mathbf{R}} &= \sum_{k=1}^{K} A_k^2 \underline{\mathbf{w}}_k \underline{\mathbf{w}}_k^2 + \sum_{k=1}^{K} A_k^2 \underline{\tilde{\mathbf{w}}}_k \underline{\tilde{\mathbf{w}}}_k^2 + \underline{\boldsymbol{\Sigma}}_v \\ &= \sum_{k=1}^{K} A_k^2 \boldsymbol{\mathcal{H}}_k \mathbf{c}_k \mathbf{c}_k^H \boldsymbol{\mathcal{H}}_k^H + \sum_{k=1}^{K} A_k^2 \tilde{\boldsymbol{\mathcal{H}}}_k \mathbf{c}_k \mathbf{c}_k^H \tilde{\boldsymbol{\mathcal{H}}}_k^H + \underline{\boldsymbol{\Sigma}}_v \\ &= \sum_{k=1}^{K} A_k^2 \underline{\boldsymbol{\mathcal{C}}}_k \mathbf{h}_k \mathbf{h}_k^H \underline{\boldsymbol{\mathcal{C}}}_k^H + \sum_{k=1}^{K} A_k^2 \underline{\tilde{\boldsymbol{\mathcal{C}}}}_k \mathbf{h}_k \mathbf{h}_k^H \underline{\tilde{\boldsymbol{\mathcal{C}}}}_k^H + \underline{\boldsymbol{\Sigma}}_v \end{aligned} \quad (1.17)$$

where

$$\underline{\boldsymbol{\Sigma}}_v = \mathrm{E}\{\underline{\mathbf{v}}(n)\underline{\mathbf{v}}(n)^H\}. \quad (1.18)$$

Introducing

$$\underline{\mathbf{W}} = [A_1 \underline{\mathbf{w}}_1, A_2 \underline{\mathbf{w}}_2, \dots, A_K \underline{\mathbf{w}}_K] \quad (1.19)$$

$$\underline{\tilde{\mathbf{W}}} = [A_1 \underline{\tilde{\mathbf{w}}}_1, A_2 \underline{\tilde{\mathbf{w}}}_2, \dots, A_K \underline{\tilde{\mathbf{w}}}_K] \quad (1.20)$$

(1.17) can be represented in a more compact form as

$$\underline{\mathbf{R}} = \underline{\mathbf{W}}\ \underline{\mathbf{W}}^H + \underline{\tilde{\mathbf{W}}}\ \underline{\tilde{\mathbf{W}}}^H + \underline{\boldsymbol{\Sigma}}_v. \quad (1.21)$$

Note that in Chapters 2-4 we only use the ISI-free part of the received signal. The ISI-free signal model of Chapter 2 directly follows from the signal model presented in this section.

Chapter 2

Signature estimation in the presence of white noise

2.1 Introduction

Among various blind signature estimation techniques, subspace-based methods [6], [47], [76], [91], [101], [111] represent a promising trend. These methods exploit the fact that user signals occupy a low-dimensional subspace in the observation space and that the signature waveform of each particular user belongs to a subspace defined by its associated spreading sequence. Although blind subspace-based signature waveform estimation techniques have a clear potential to provide excellent estimation accuracy, certain practical hurdles may make them inapplicable to real-world environments. In particular, an essential shortcoming of the subspace-based techniques is that they are only applicable to underloaded systems, i.e., systems wherein the number of active users is less than the dimension of the observation space [6], [47], [76]. Another practical concern emerges from the fact that the signature waveform of each user represents the convolution of the associated spreading sequence with the channel impulse response. Hence, the problem of signature waveform estimation boils down to the channel identification problem [6], [47], [76], [89], [91], [111]. Since all of the aforementioned subspace-based methods implicitly [6], [89], [91] or explicitly [47], [76], [111] pose some restrictions on the length of the channel, their reliability in a long-delay multi-path environment is questionable. The major contribution of this chapter (see also [109],

[111]) is to propose an improved version of one popular subspace-based signature waveform estimation technique which, as compared to its conventional counterpart, is capable to identify longer channels in more heavily loaded environments.

While the idea of the proposed technique can be easily adapted to modify other subspace-based techniques, we develop our method based on the popular algorithm by Liu and Xu [47], which is hereafter referred to as the LX algorithm. To develop our approach, we assume that all transmitted symbols are drawn from the BPSK constellation and the ambient noise is a white circular process [30], [60]. We exploit the noise circularity property jointly with the non-circular property of the transmitted signals to effectively double the dimension of the observation space while keeping the dimension of the signal subspace unchanged. Using such an extended observation space, the proposed technique is shown to facilitate identification of substantially longer channels in more heavily loaded environments than the LX algorithm. Moreover, as the extended model developed below provides more equations to estimate the sampled channel impulse response, our technique is shown to substantially outperform the LX algorithm when both techniques are implemented using the LS method.

We prove necessary and sufficient conditions for channel identifiability of the proposed technique and show that this condition is *weaker* than that of the LX algorithm. This means that any channel which could be identified using the LX algorithm is also identifiable by our technique, while the reverse statement is not necessarily true. Using the first-order perturbation theory, we also derive the closed-form expressions for the MSE of the proposed estimation algorithm. These expressions explicitly clarify how the performance of the algorithm depends on environmental parameters.

It should be stressed that the non-circularity property of the BPSK signals was already used in direction finding [13], [27] and multiuser detection [7],[10], [22], [24], [68], [69], [81], [104]. However, the signature waveform estimation problem has not been addressed in any of these papers. It is also worth noting that, for any subspace-based signature waveform estimation algorithm, it is possible to increase the observation space by means of oversampling the received data in time or in space (in the latter case, multiple receive antennas have to be used) [47]. In particular, temporal or spatial oversampling by the

factor of two doubles the dimension of the observation space. However, such oversampling may result in temporally or spatially correlated noise [2], [66] which makes the conventional subspace techniques inapplicable. Additionally, it can substantially increase the receiver cost and necessitate extra-hardware. Nevertheless, even if it is possible to take an advantage of the oversampling approach, our algorithm can further enlarge the observation space on top of the improvements achieved by oversampling. Temporal oversampled extension of both the LX algorithm and the proposed technique are also briefly presented.

The rest of this chapter is organized as follows. In Section 2.2, the signal model is formulated. Section 2.3 provides an overview of the LX algorithm. Section 2.4 presents the proposed method. Simulation results are given in Section 2.5, and Section 2.6 briefly summarizes the chapter.

2.2 ISI-free signal model

Note from assumption **A2** that, the duration of the user channels are much shorter than the symbol period T_s, and the effect of ISI can be neglected [32], [47]. Throughout Chapters 2-4 we use the latter result and consider only the ISI-free part of the received signal. From (1.7) it follows that the $(L_c - L + 1) \times 1$ ISI-free received sampled data vector can be written as

$$\mathbf{x}(n) = \sum_{k=1}^{K} A_k b_k(n) \mathbf{w}_k + \mathbf{v}(n) \tag{2.1}$$

where

$$\mathbf{x}(n) = [x(nT_s + (L-1)T_c), x(nT_s + LT_c), \ldots, x(nT_s + (L_c-1)T_c)]^T \tag{2.2}$$

$$\mathbf{w}_k = [w_k((L-1)T_c), w_k(LT_c), \ldots, w_k((L_c-1)T_c)]^T \tag{2.3}$$

$$\mathbf{v}(n) = [v(nT_s + (L-1)T_c), v(nT_s + LT_c), \ldots, v(nT_s + (L_c-1)T_c)]^T \tag{2.4}$$

and

$$\mathbf{w}_k = \begin{bmatrix} c_k[L-1] & \ldots & c_k[0] \\ c_k[L] & \ldots & c_k[1] \\ \vdots & \ddots & \vdots \\ c_k[L_c-1] & \ldots & c_k[L_c-L] \end{bmatrix} \mathbf{h}_k = \boldsymbol{\mathcal{C}}_k \mathbf{h}_k. \tag{2.5}$$

Since the spreading code of the desired user is known at the receiver, once the channel vector $\mathbf{h}_k$ is estimated, $\mathbf{w}_k$ can be directly obtained from (2.5) [47]. Hence, unless otherwise stated, throughout this thesis we consider the problem of channel vector estimation rather than the signature vector estimation. For the sake of consistency with the paper [47] and without any loss of generality, we assume in Chapters 2-4 that $\mathbf{h}_k$ is a unit-norm vector ($\|\mathbf{h}_k\| = 1$), i.e., the norm of $\mathbf{h}_k$ is absorbed in the corresponding amplitude A_k. One can rewrite (2.1) in a more compact form as [47]

$$\mathbf{x}(n) = \mathbf{W}\mathbf{b}(n) + \mathbf{v}(n) \tag{2.6}$$

where

$$\mathbf{W} = [A_1\mathbf{w}_1, A_2\mathbf{w}_2, \ldots, A_K\mathbf{w}_K] \tag{2.7}$$

is assumed to be a full-rank matrix, and

$$\mathbf{b}(n) = [b_1(n), b_2(n), \ldots, b_K(n)]^T . \tag{2.8}$$

2.3 The LX algorithm

The covariance matrix of the ISI-free part of data is given by

$$\mathbf{R} = \mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}(n)^H\right\} . \tag{2.9}$$

From (2.6) and the assumption that noise is a white random process we have [47]

$$\mathbf{R} = \mathbf{W}\mathbf{W}^H + \sigma^2\mathbf{I} \tag{2.10}$$

where $\sigma^2\mathbf{I}$ is the noise covariance matrix and

$$\sigma^2 = \mathrm{E}\left\{|v(t)|^2\right\} . \tag{2.11}$$

The matrix (2.10) can be eigendecomposed as

$$\mathbf{R} = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Omega}_s & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \tag{2.12}$$

where the $K \times K$ diagonal matrix

$$\mathbf{\Omega}_s = \mathbf{\Lambda}_s + \sigma^2 \mathbf{I} = \text{diag}(\lambda_1 + \sigma^2, \ldots, \lambda_K + \sigma^2) \tag{2.13}$$

contains the K largest (signal-subspace) eigenvalues of $\mathbf{R}$, the $(L_c - L + 1) \times K$ matrix $\mathbf{U}_s$ contains the corresponding orthonormal eigenvectors, and $\{\lambda_k\}_{k=1}^{K}$ are the non-zero eigenvalues of the matrix $\mathbf{W}\mathbf{W}^H$. In turn, the $(L_c - L + 1) \times (L_c - L - K + 1)$ matrix $\mathbf{U}_n$ contains the orthonormal noise-subspace eigenvectors of $\mathbf{R}$ associated with the multiple eigenvalue σ^2. Without any loss of generality, we assume that the first user is the user-of-interest and, therefore, $\mathbf{h}_1$ is the channel vector to be identified. As

$$\texttt{range}(\mathbf{W}) = \texttt{range}(\mathbf{U}_s) \tag{2.14}$$

we have [47]

$$\mathbf{U}_n^H \mathbf{w}_1 = \mathbf{0}. \tag{2.15}$$

It follows from (2.15) that $\mathbf{h}_1$ is a nontrivial solution of the following equation:

$$\mathbf{T}\mathbf{h} = \mathbf{0} \tag{2.16}$$

where

$$\mathbf{T} = \mathbf{U}_n^H \boldsymbol{\mathcal{C}}_1 \tag{2.17}$$

is an $L_c - L + 1 - K \times L$ matrix. As (2.16) is a set of $L_c - L - K + 1$ complex equations with L complex unknowns, the uniqueness of its nontrivial solution requires that (2.16) is not under-determined, or, equivalently [47]

$$2L + K \leq L_c + 1. \tag{2.18}$$

Equation (2.18) represents the necessary condition for unique nontrivial solution of (2.16). It has been shown in [47] that, up to an arbitrary scaling factor, $\mathbf{h}_1$ can be uniquely identified from (2.16) if and only if $\boldsymbol{\mathcal{C}}_1$ is a full column-rank matrix and

$$\texttt{dim}\,\{\texttt{range}(\boldsymbol{\mathcal{C}}_1) \cap \texttt{range}(\mathbf{W})\} = 1. \tag{2.19}$$

Note that if (2.19) holds but $\boldsymbol{\mathcal{C}}_1$ is not a full-column rank matrix, then there is no guarantee that $\mathbf{h}_1$ can be uniquely identified from (2.16). In spite of this fact, $\mathbf{w}_1$ is still uniquely identifiable up to an arbitrary scaling factor as $\boldsymbol{\mathcal{C}}_1\tilde{\mathbf{h}}$ where $\tilde{\mathbf{h}}$ is a solution to (2.16) [47].

In practice, $\mathbf{R}$ is not known exactly and can be estimated as

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}(n)^H. \tag{2.20}$$

As a result, $\mathbf{U}_n$ is estimated as $\hat{\mathbf{U}}_n$ that consists of the eigenvectors associated with the smallest $L_c - L + 1 - K$ eigenvalues of $\hat{\mathbf{R}}$. Substituting $\hat{\mathbf{U}}_n$ in lieu of $\mathbf{U}_n$ in (2.15) and solving the so-obtained equation in the LS sense, we have that the estimated channel vector $\hat{\mathbf{h}}_1$ is given by [47]

$$\hat{\mathbf{h}}_1 = \mho\left(\boldsymbol{\mathcal{C}}_1^H\hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^H\boldsymbol{\mathcal{C}}_1\right). \tag{2.21}$$

Using the first-order perturbation theory [40], [99], the encountered mean-squared estimation error $\delta\mathbf{h}_1 = \hat{\mathbf{h}}_1 - \mathbf{h}_1$ can be approximately written as [99]

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2}{N}\|\mathbf{T}^\dagger\|_{\mathrm{F}}^2\mathbf{w}_1^H\big(\mathbf{U}_s\boldsymbol{\Omega}_s^{-1}\mathbf{U}_s^H + \sigma^2\mathbf{U}_s\boldsymbol{\Omega}_s^{-2}\mathbf{U}_s^H\big)\mathbf{w}_1. \tag{2.22}$$

Assuming that the signatures of different users are orthogonal to each other, i.e.,

$$\mathbf{w}_i^H\mathbf{w}_j = \|\mathbf{w}_i\|^2\delta_{ij} \tag{2.23}$$

the MSE expression (2.22) can be significantly simplified. Note that due to multipath effects, the orthogonality assumption of the signature vectors does not perfectly hold in practice. However, CDMA spreading codes are deliberately designed so that, even after passing through a frequency selective channel, the resulting spread-spectrum signature waveforms occupy a wide frequency band and behave as almost white pseudo-random signals [3], [8], [48], [61], [82]. Hence, in most practical scenarios (2.23) approximately holds as the signature vectors are sampled versions of almost white pseudo-random signature waveforms.

It directly follows from (2.23) that

$$\mathbf{U}_s = \left[\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \ldots, \frac{\mathbf{w}_K}{\|\mathbf{w}_K\|}\right] \tag{2.24}$$

$$\mathbf{\Omega}_s = \mathrm{diag}\left(A_1^2\|\mathbf{w}_1\|^2, A_2^2\|\mathbf{w}_2\|^2, \ldots, A_K^2\|\mathbf{w}_K\|^2\right). \tag{2.25}$$

Substituting (2.24) and (2.25) into (2.22) and using (2.23) yields

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2\|\mathbf{T}^\dagger\|_{\mathrm{F}}^2}{NA_1^2}\left(1 + \frac{\sigma^2}{A_1^2\|\mathbf{w}_1\|^2}\right). \tag{2.26}$$

If SNR is high enough, i.e.,

$$\sigma^2 \ll A_1^2\|\mathbf{w}_1\|^2 \tag{2.27}$$

then, the MSE of the channel estimate can be further simplified to

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2\|\mathbf{T}^\dagger\|_{\mathrm{F}}^2}{NA_1^2}. \tag{2.28}$$

An expression equivalent to (2.28) has been derived for the MSE of the estimated signature, $\boldsymbol{\mathcal{C}}_1\hat{\mathbf{h}}_1$, in [47]. The accuracy of (2.26) is validated in Section 2.5.

2.3.1 Temporally oversampled version of the LX algorithm

Condition (2.18) severely restricts both the maximum channel order and the number of active users. The maximum values of L and K can be increased if more equations are available to identify the channel vector. This could be achieved by means of temporal oversampling [47], [90]. Let us sample the continuous-time received signal (1.1) with the rate κ/T_c in the interval corresponding to the nth transmitted symbol where the oversampling factor $\kappa \geq 1$ is an integer. The ISI-free part of the oversampled data vector is given by [90]

$$\mathbf{x}_\kappa(n) = \sum_{k=1}^{K} A_k b_k(n)\mathbf{w}_{k\kappa} + \mathbf{v}_\kappa(n) \tag{2.29}$$

where

$$\mathbf{x}_\kappa(n) = [\mathbf{x}\{n, L-1\}^T, \mathbf{x}\{n, L\}^T, \ldots, \mathbf{x}\{n, L_c-1\}^T]^T \tag{2.30}$$

$$\mathbf{w}_{k\kappa} = [\mathbf{w}_k\{L-1\}^T, \mathbf{w}_k\{L\}^T, \ldots, \mathbf{w}_k\{L_c-1\}^T]^T \tag{2.31}$$

$$\mathbf{v}_\kappa(n) = [\mathbf{v}\{n, L-1\}^T, \mathbf{v}\{n, L\}^T, \ldots, \mathbf{v}\{n, L_c-1\}^T]^T \tag{2.32}$$

with

$$\mathbf{x}\{n,i\} = \left[x(nT_s+iT_c), x\left(nT_s+\left(i+\frac{1}{\kappa}\right)T_c\right),\ldots,x\left(nT_s+\left(i+\frac{\kappa-1}{\kappa}\right)T_c\right)\right]^T \quad (2.33)$$

$$\mathbf{w}_k\{i\} = \left[w_k(iT_c), w_k\left(\left(i+\frac{1}{\kappa}\right)T_c\right),\ldots,w_k\left(\left(i+\frac{\kappa-1}{\kappa}\right)T_c\right)\right]^T \quad (2.34)$$

$$\mathbf{v}\{n,i\} = \left[v(nT_s+iT_c), v\left(nT_s+\left(i+\frac{1}{\kappa}\right)T_c\right),\ldots,v\left(nT_s+\left(i+\frac{\kappa-1}{\kappa}\right)T_c\right)\right]^T \quad (2.35)$$

Using (1.3), (2.31), and (2.34), it follows that

$$\mathbf{w}_{k\kappa} = \boldsymbol{\mathcal{C}}_{k\kappa}\mathbf{h}_{k\kappa} \quad (2.36)$$

where

$$\boldsymbol{\mathcal{C}}_{k\kappa} = \boldsymbol{\mathcal{C}}_k \otimes \mathbf{I}_\kappa \quad (2.37)$$

$$\mathbf{h}_{k\kappa} = [\mathbf{h}_k\{0\}^T, \mathbf{h}_k\{1\}^T,\ldots,\mathbf{h}_k\{L-1\}^T]^T \quad (2.38)$$

$$\mathbf{h}_k\{i\} = \left[h_k(iT_c), h_k\left(\left(i+\frac{1}{\kappa}\right)T_c\right),\ldots,h_k\left(\left(i+\frac{\kappa-1}{\kappa}\right)T_c\right)\right]^T. \quad (2.39)$$

Let us write (2.29) as

$$\mathbf{x}_\kappa(n) = \mathbf{W}_\kappa\mathbf{b}(n) + \mathbf{v}_\kappa(n) \quad (2.40)$$

where

$$\mathbf{W}_\kappa = [A_1\mathbf{w}_{1\kappa}, A_2\mathbf{w}_{2\kappa},\ldots,A_K\mathbf{w}_{K\kappa}]. \quad (2.41)$$

From assumption **A3** and (2.40) we have that

$$\mathbf{R}_\kappa = \mathrm{E}\left\{\mathbf{x}_\kappa(n)\mathbf{x}_\kappa(n)^H\right\} = \mathbf{W}_\kappa\mathbf{W}_\kappa^H + \sigma^2\mathbf{I}. \quad (2.42)$$

Note that $\mathbf{R}_\kappa$ is a $\kappa(L_c-L+1)\times\kappa(L_c-L+1)$ matrix and

$$\mathtt{dim}\{\mathtt{range}(\mathbf{W}_\kappa)\} = K. \quad (2.43)$$

Similar to our discussion prior to (2.16), it can be shown that $\mathbf{h}_{1\kappa}$ is a solution to

$$\mathbf{T}_\kappa\mathbf{h} = \mathbf{0} \quad (2.44)$$

where

$$\mathbf{T}_\kappa = \mathbf{U}_{n_\kappa}^H\boldsymbol{\mathcal{C}}_{1\kappa} \quad (2.45)$$

and $\mathbf{U}_{n_\kappa}$ contains the $\kappa(L_c - L + 1) - K$ orthonormal noise-subspace eigenvectors of $\mathbf{R}_\kappa$ associated with the multiple eigenvalue σ^2. As (2.44) has $\kappa(L_c - L + 1) - K$ complex equations and κL complex unknowns, to have a unique nontrivial solution to (2.44), it is necessary that

$$2\kappa L + K \leq \kappa(L_c + 1). \tag{2.46}$$

Comparing (2.18) with (2.46), we have that using the oversampling factor of κ, the maximum admissible number of users is increased by κ times. It can also be observed from (2.46) that for a fixed number of active users, the oversampling technique considerably increases the maximum admissible channel length. It should be stressed that if multiple receive antennas are available, similar improvements could be achieved by means of spatial oversampling [47].

2.4 The proposed algorithm for BPSK-modulated systems

As discussed in the previous section, one could use the temporal oversampling technique to extend the dimension of the observation space. Then, the oversampled data can be used in the LX algorithm to facilitate estimating longer channels in more heavily loaded environments. However, oversampling may result in a correlated noise [2], [66] which makes the conventional subspace-based estimation techniques (including the LX algorithm) inapplicable.

As shown below, if the transmitted symbols are non-circular, then it is possible to extend the observation space without temporal oversampling. Let us assume that the transmitted symbols are drawn from the BPSK constellation, and, therefore, they are non-circular, while the white noise is circular. Due to the circular property of noise, we have [30], [60]

$$\mathrm{E}\{\mathbf{v}(n)\mathbf{v}(n)^T\} = \mathbf{0}. \tag{2.47}$$

Introducing the $2(L_c - L + 1) \times 1$ extended received data vectors as

$$\bar{\mathbf{x}}(n) = [\mathbf{x}^T(n)\ \mathbf{x}^H(n)]^T \tag{2.48}$$

we have that

$$\bar{\mathbf{x}}(n) = \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \mathbf{b}(n) + \begin{bmatrix} \mathbf{v}(n) \\ \mathbf{v}^*(n) \end{bmatrix}. \tag{2.49}$$

Using equation (2.47) along with equation (2.49) yields

$$\bar{\mathbf{R}} = \mathrm{E}\{\bar{\mathbf{x}}(n)\bar{\mathbf{x}}^H(n)\} = \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H + \sigma^2\mathbf{I}. \tag{2.50}$$

From the fact that $\mathbf{W}$ is a full column rank matrix, it can be concluded that $[\mathbf{W}^T \;\; \mathbf{W}^H]^T$ is also full column rank. We can show this by contradiction as follows. Assume that $[\mathbf{W}^T \;\; \mathbf{W}^H]^T$ is not full column-rank. Then, there exists a vector $\mathbf{g} \neq \mathbf{0}$ such that

$$\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \mathbf{g} = \mathbf{0} \tag{2.51}$$

but (2.51) also implies that $\mathbf{W}\mathbf{g} = \mathbf{0}$. This, however, contradicts to the fact that $\mathbf{W}$ is full column rank.

The eigendecomposition of $\bar{\mathbf{R}}$ can be written as

$$\bar{\mathbf{R}} = \begin{bmatrix} \bar{\mathbf{U}}_s & \bar{\mathbf{U}}_n \end{bmatrix} \begin{bmatrix} \bar{\mathbf{\Omega}}_s & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{U}}_s^H \\ \bar{\mathbf{U}}_n^H \end{bmatrix} \tag{2.52}$$

where the $K \times K$ matrix

$$\bar{\mathbf{\Omega}}_s = \bar{\mathbf{\Lambda}}_s + \sigma^2\mathbf{I} = \mathrm{diag}(\bar{\lambda}_1 + \sigma^2, \ldots, \bar{\lambda}_K + \sigma^2) \tag{2.53}$$

contains the K largest eigenvalues of $\bar{\mathbf{R}}$, the $2(L_c - L + 1) \times K$ matrix $\bar{\mathbf{U}}_s$ contains the corresponding orthonormal (signal-subspace) eigenvectors, the $2(L_c - L + 1) \times (2(L_c - L + 1) - K)$ matrix $\bar{\mathbf{U}}_n$ is built from the orthonormal noise-subspace eigenvectors of $\bar{\mathbf{R}}$ that correspond to the eigenvalue σ^2 of this matrix, and $\{\bar{\lambda}_k\}_{k=1}^K$ are the non-zero (signal-subspace) eigenvalues of the matrix $\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H$. Obviously, using $\bar{\mathbf{x}}$ instead of $\mathbf{x}$, the dimension of the observation space is increased twice, while the dimension of the signal subspace remains unchanged. Since

$$\mathtt{range}(\bar{\mathbf{U}}_s) = \mathtt{range}\left(\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}\right) \tag{2.54}$$

we have

$$\bar{\mathbf{U}}_n^H \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} = \mathbf{0}. \tag{2.55}$$

From (2.55), it follows that

$$\bar{\mathbf{U}}_n^H \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_1^* \end{bmatrix} = \bar{\mathbf{U}}_n^H \begin{bmatrix} \boldsymbol{\mathcal{C}}_1 \mathbf{h}_1 \\ \boldsymbol{\mathcal{C}}_1^* \mathbf{h}_1^* \end{bmatrix} = \mathbf{0}. \tag{2.56}$$

Denoting the matrix which contains the first $L_c - L + 1$ rows of $\bar{\mathbf{U}}_n$ as $\bar{\mathbf{U}}_{n_1}$, and the matrix which contains the remaining last $L_c - L + 1$ rows of $\bar{\mathbf{U}}_n$ as $\bar{\mathbf{U}}_{n_2}$, we obtain from (2.56) that

$$\bar{\mathbf{U}}_{n_1}^H \boldsymbol{\mathcal{C}}_1 \mathbf{h}_1 + \bar{\mathbf{U}}_{n_2}^H \boldsymbol{\mathcal{C}}_1^* \mathbf{h}_1^* = \mathbf{0}. \tag{2.57}$$

Equation (2.57) can be solved for $\mathbf{h}_1$ as follows. Let

$$\bar{\mathbf{T}}_1 = \bar{\mathbf{U}}_{n_1}^H \boldsymbol{\mathcal{C}}_1 \tag{2.58}$$

$$\bar{\mathbf{T}}_2 = \bar{\mathbf{U}}_{n_2}^H \boldsymbol{\mathcal{C}}_1^* \tag{2.59}$$

and let us rewrite (2.57) in the following equivalent form

$$\begin{bmatrix} \mathrm{Re}(\bar{\mathbf{T}}_1) + \mathrm{Re}(\bar{\mathbf{T}}_2) & \mathrm{Im}(\bar{\mathbf{T}}_2) - \mathrm{Im}(\bar{\mathbf{T}}_1) \\ \mathrm{Im}(\bar{\mathbf{T}}_1) + \mathrm{Im}(\bar{\mathbf{T}}_2) & \mathrm{Re}(\bar{\mathbf{T}}_1) - \mathrm{Re}(\bar{\mathbf{T}}_2) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\mathbf{h}_1) \\ \mathrm{Im}(\mathbf{h}_1) \end{bmatrix} = \mathbf{0}. \tag{2.60}$$

As an alternative of using the LX algorithm, we propose to use (2.60) for signature waveform estimation. In the finite sample case, our technique can be summarized as the following sequence of steps.

1. Find the sample estimate

$$\hat{\bar{\mathbf{R}}} = \frac{1}{N} \sum_{k=1}^{N} \bar{\mathbf{x}}(k) \bar{\mathbf{x}}^H(k) \tag{2.61}$$

   of the covariance matrix $\bar{\mathbf{R}}$.

2. Compute the eigendecomposition of $\hat{\bar{\mathbf{R}}}$ in the form

$$\hat{\bar{\mathbf{R}}} = \begin{bmatrix} \hat{\bar{\mathbf{U}}}_s & \hat{\bar{\mathbf{U}}}_n \end{bmatrix} \begin{bmatrix} \hat{\bar{\mathbf{\Omega}}}_s & \mathbf{0} \\ \mathbf{0} & \hat{\bar{\mathbf{\Omega}}}_n \end{bmatrix} \begin{bmatrix} \hat{\bar{\mathbf{U}}}_s^H \\ \hat{\bar{\mathbf{U}}}_n^H \end{bmatrix} \tag{2.62}$$

where the matrices $\hat{\bar{\mathbf{U}}}_s$, $\hat{\bar{\mathbf{\Omega}}}_s$ and $\hat{\bar{\mathbf{U}}}_n$ are the finite-sample counterparts of the matrices $\bar{\mathbf{U}}_s$, $\bar{\mathbf{\Omega}}_s$ and $\bar{\mathbf{U}}_n$, respectively, and $\hat{\bar{\mathbf{\Omega}}}_n$ is the $(2(L_c-L+1)-K)\times(2(L_c-L+1)-K)$ matrix of the noise-subspace eigenvalues of $\hat{\bar{\mathbf{R}}}$.

3. Using the first and the last (L_c-L+1) rows of $\hat{\bar{\mathbf{U}}}_n$, obtain the matrices $\hat{\bar{\mathbf{U}}}_{n_1}$ and $\hat{\bar{\mathbf{U}}}_{n_2}$ which are the finite-sample counterparts of the matrices $\bar{\mathbf{U}}_{n_1}$ and $\bar{\mathbf{U}}_{n_2}$, respectively. Compute

$$\hat{\bar{\mathbf{T}}}_1 = \hat{\bar{\mathbf{U}}}_{n_1}^H \boldsymbol{\mathcal{C}}_1 \tag{2.63}$$

$$\hat{\bar{\mathbf{T}}}_2 = \hat{\bar{\mathbf{U}}}_{n_2}^H \boldsymbol{\mathcal{C}}_1^*. \tag{2.64}$$

4. Compute the matrix

$$\mathbf{\Psi} = \begin{bmatrix} \mathrm{Re}(\hat{\bar{\mathbf{T}}}_1)+\mathrm{Re}(\hat{\bar{\mathbf{T}}}_2) & \mathrm{Im}(\hat{\bar{\mathbf{T}}}_2)-\mathrm{Im}(\hat{\bar{\mathbf{T}}}_1) \\ \mathrm{Im}(\hat{\bar{\mathbf{T}}}_1)+\mathrm{Im}(\hat{\bar{\mathbf{T}}}_2) & \mathrm{Re}(\hat{\bar{\mathbf{T}}}_1)-\mathrm{Re}(\hat{\bar{\mathbf{T}}}_2) \end{bmatrix} \tag{2.65}$$

and find the $2L\times 1$ vector

$$\mathbf{s} = \mathcal{U}(\mathbf{\Psi}^H\mathbf{\Psi}). \tag{2.66}$$

5. Estimate the channel vector $\mathbf{h}_1$ as

$$\hat{\mathbf{h}}_1 = \mathbf{s}_1 + j\mathbf{s}_2 \tag{2.67}$$

where $\mathbf{s}_1$ and $\mathbf{s}_2$ are the $L\times 1$ subvectors of $\mathbf{s}$ which contain its first and last L entries, respectively.

The linear system in (2.60) contains $4(L_c-L+1)-2K$ real equations and $2L$ real unknown variables. To have a unique nontrivial solution for (2.60), it is necessary that the number of unknowns is less than or equal to the number of equations, that is,

$$3L+K \le 2(L_c+1). \tag{2.68}$$

Comparing (2.68) with (2.18) and taking into account that L_c is much larger than L, it can be observed that the proposed algorithm has a substantially more relaxed condition

on the maximal identifiable channel length L and the number of active users K than the conventional LX algorithm. Moreover, the fact that the number of real equations is increased in the proposed algorithm from $2(L_c - L + 1) - 2K$ to $4(L_c - L + 1) - 2K$ as compared to the LX algorithm (while keeping the number of unknowns unchanged) implies that the channel estimation accuracy of our algorithm is expected to improve with respect to the LX technique.

2.4.1 Identifiability condition

In this section, we derive the necessary and sufficient non-identifiability condition of the proposed algorithm. Using of this condition, we will show that any channel vector which can be identified from the LX algorithm is always identifiable using our approach, while the reverse statement in not necessarily true.

First of all, note that

$$\boldsymbol{\mathcal{C}}_1\mathbf{h}_1 = \mathbf{w}_1 = \mathbf{W}\mathbf{e}_1/A_1. \tag{2.69}$$

Hence,

$$\mathtt{dim}\{\mathtt{range}(\boldsymbol{\mathcal{C}}_1) \cap \mathtt{range}(\mathbf{W})\} \geq 1. \tag{2.70}$$

The following theorem establishes the identifiability result for our technique.

Theorem 1: Let $\boldsymbol{\mathcal{C}}_1$ and $\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}$ be full column-rank matrices and ξ be an arbitrary scalar. Then, the equation

$$\bar{\mathbf{U}}_{n_1}^H \boldsymbol{\mathcal{C}}_1\mathbf{h} + \bar{\mathbf{U}}_{n_2}^H \boldsymbol{\mathcal{C}}_1^*\mathbf{h}^* = \mathbf{0} \tag{2.71}$$

has a nontrivial solution other than $\xi\mathbf{h}_1$ if and only if the following two conditions jointly hold:

C-1 There exist two vectors $\tilde{\mathbf{h}} \neq \xi\mathbf{h}_1$ and $\mathbf{f}$ such that

$$\boldsymbol{\mathcal{C}}_1\tilde{\mathbf{h}} = \mathbf{W}\mathbf{f} \tag{2.72}$$

which is equivalent to

$$\mathtt{dim}\{\mathtt{range}(\boldsymbol{\mathcal{C}}_1) \cap \mathtt{range}(\mathbf{W})\} > 1. \tag{2.73}$$

C-2 For any vector $\mathbf{f}$ satisfying **C-1**,

$$\mathbf{W}(\mathbf{f} - \mathbf{f}^*) = \mathbf{0}. \tag{2.74}$$

Proof: Equation (2.71) can be rewritten as

$$\bar{\mathbf{U}}_n^H \begin{bmatrix} \boldsymbol{\mathcal{C}}_1 \mathbf{h} \\ \boldsymbol{\mathcal{C}}_1^* \mathbf{h}^* \end{bmatrix} = \mathbf{0}. \tag{2.75}$$

We start our proof with the necessity part. Let (2.75) be satisfied with some $\tilde{\mathbf{h}} \neq \xi \mathbf{h}_1$. As the matrix $\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}$ is full column-rank, this directly yields

$$\begin{bmatrix} \boldsymbol{\mathcal{C}}_1 \tilde{\mathbf{h}} \\ \boldsymbol{\mathcal{C}}_1^* \tilde{\mathbf{h}}^* \end{bmatrix} \in \mathtt{range}\left(\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \right). \tag{2.76}$$

From (2.76), it follows that there exists a vector $\mathbf{f}$ such that

$$\begin{bmatrix} \boldsymbol{\mathcal{C}}_1 \tilde{\mathbf{h}} \\ \boldsymbol{\mathcal{C}}_1^* \tilde{\mathbf{h}}^* \end{bmatrix} = \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \mathbf{f}. \tag{2.77}$$

Hence, we have (2.72). Then, (2.73) follows from (2.72) along with (2.69) and the fact that the vectors $\mathbf{h}_1$ and $\tilde{\mathbf{h}}$ are linearly independent. This proves **C-1**. To prove **C-2**, we note that from (2.77)

$$\mathbf{W}^* \mathbf{f} = \boldsymbol{\mathcal{C}}_1^* \tilde{\mathbf{h}}^* = (\boldsymbol{\mathcal{C}}_1 \tilde{\mathbf{h}})^* = \mathbf{W}^* \mathbf{f}^* \tag{2.78}$$

which directly results in **C-2**.

To prove the sufficiency part, let us assume that both **C-1** and **C-2** hold. It follows from (2.74) that $\mathbf{W}^* \mathbf{f}^* = \mathbf{W}^* \mathbf{f}$. From the latter result and (2.72), we have

$$\boldsymbol{\mathcal{C}}_1^* \tilde{\mathbf{h}}^* = \mathbf{W}^* \mathbf{f}. \tag{2.79}$$

Using (2.79) along with (2.72), we obtain (2.77), or, equivalently, (2.76). From (2.76) it follows that $\tilde{\mathbf{h}} \neq \xi \mathbf{h}_1$ is a solution to equation (2.75) (or, equivalently, (2.71)). This completes the proof. □

It should be noted that, according to (2.19), **C-1** by itself is the necessary and sufficient condition for the channel non-identifiability of the LX algorithm (which is, obviously, *weaker* than that of the proposed algorithm). This implies that the identifiability condition of the LX algorithm is *stronger* than that of the proposed algorithm. Therefore, any channel vector identifiable by means of the LX algorithm is also identifiable using the proposed technique while the reverse statement does not necessarily hold.

2.4.2 Performance analysis

In practice, the exact covariance matrix $\bar{\mathbf{R}}$ is unavailable. In this section, we use the first-order perturbation theory to derive the MSE of the channel vector estimate obtained by the algorithm of Section 2.4 where the sample estimate of $\bar{\mathbf{R}}$ is used.

Theorem 2: Assume that the channel vector $\mathbf{h}_1$ is estimated using the algorithm of Section 2.4. Then, the first-order perturbation theory-based approximation of the MSE of the estimated channel vector is given by

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2\|\bar{\mathbf{T}}^\dagger\|_F^2}{2N} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_1^* \end{bmatrix}^H \bar{\mathbf{U}}_s\bar{\mathbf{\Lambda}}_s^{-1}\bar{\mathbf{\Omega}}_s\bar{\mathbf{\Lambda}}_s^{-1}\bar{\mathbf{U}}_s^H \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_1^* \end{bmatrix} \tag{2.80}$$

where

$$\delta\mathbf{h}_1 = \hat{\mathbf{h}}_1 - \mathbf{h}_1 \tag{2.81}$$

and

$$\bar{\mathbf{T}} = \left[\bar{\mathbf{T}}_1 \;\; \bar{\mathbf{T}}_2\right]. \tag{2.82}$$

Moreover, if (2.23) and (2.27) hold, then (2.80) reduces to

$$\mathrm{E}\{\|\delta\mathbf{h}_1\|^2\} \approx \frac{\sigma^2\|\bar{\mathbf{T}}^\dagger\|_F^2}{2A_1^2N}. \tag{2.83}$$

Proof: Let us introduce

$$\delta\bar{\mathbf{R}} = \hat{\bar{\mathbf{R}}} - \bar{\mathbf{R}} \tag{2.84}$$

$$\delta\bar{\mathbf{U}}_n = \hat{\bar{\mathbf{U}}}_n - \bar{\mathbf{U}}_n. \tag{2.85}$$

Using the perturbation theory [40], [99], the first-order estimate of $\delta\bar{\mathbf{U}}_n$ can be written as

$$\delta\bar{\mathbf{U}}_n \doteq -\bar{\mathbf{U}}_s\bar{\mathbf{\Lambda}}_s^{-1}\bar{\mathbf{U}}_s^H\delta\bar{\mathbf{R}}\bar{\mathbf{U}}_n. \tag{2.86}$$

Since

$$\left[\begin{array}{cc}\hat{\bar{\mathbf{U}}}_{n_1}^H\boldsymbol{\mathcal{C}}_1 & \hat{\bar{\mathbf{U}}}_{n_2}^H\boldsymbol{\mathcal{C}}_1^*\end{array}\right]\left[\begin{array}{c}\hat{\mathbf{h}}_1\\ \hat{\mathbf{h}}_1^*\end{array}\right] = \left(\bar{\mathbf{T}} + \left[\begin{array}{cc}\delta\bar{\mathbf{U}}_{n_1}^H\boldsymbol{\mathcal{C}}_1 & \delta\bar{\mathbf{U}}_{n_2}^H\boldsymbol{\mathcal{C}}_1^*\end{array}\right]\right)\left(\left[\begin{array}{c}\mathbf{h}_1\\ \mathbf{h}_1^*\end{array}\right] + \left[\begin{array}{c}\delta\mathbf{h}_1\\ \delta\mathbf{h}_1^*\end{array}\right]\right) \approx \mathbf{0} \tag{2.87}$$

it follows that

$$\left[\begin{array}{c}\delta\mathbf{h}_1\\ \delta\mathbf{h}_1^*\end{array}\right] \approx -\bar{\mathbf{T}}^\dagger\left[\begin{array}{cc}\delta\bar{\mathbf{U}}_{n_1}^H\boldsymbol{\mathcal{C}}_1 & \delta\bar{\mathbf{U}}_{n_2}^H\boldsymbol{\mathcal{C}}_1^*\end{array}\right]\left[\begin{array}{c}\mathbf{h}_1\\ \mathbf{h}_1^*\end{array}\right] \tag{2.88}$$

or, equivalently,

$$\left[\begin{array}{c}\delta\mathbf{h}_1\\ \delta\mathbf{h}_1^*\end{array}\right] \approx -\bar{\mathbf{T}}^\dagger\delta\bar{\mathbf{U}}_n^H\left[\begin{array}{c}\mathbf{w}_1\\ \mathbf{w}_1^*\end{array}\right]. \tag{2.89}$$

Substituting (2.86) into the right-hand side of (2.89) and applying the expectation operation to the squared norm of the resulting expression, we obtain

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{1}{2}\left[\begin{array}{c}\mathbf{w}_1\\ \mathbf{w}_1^*\end{array}\right]^H\bar{\mathbf{U}}_s\bar{\mathbf{\Lambda}}_s^{-1}\bar{\mathbf{U}}_s^H\mathrm{E}\left\{\delta\bar{\mathbf{R}}\mathbf{\Phi}\delta\bar{\mathbf{R}}\right\}\bar{\mathbf{U}}_s\bar{\mathbf{\Lambda}}_s^{-1}\bar{\mathbf{U}}_s^H\left[\begin{array}{c}\mathbf{w}_1\\ \mathbf{w}_1^*\end{array}\right] \tag{2.90}$$

where

$$\mathbf{\Phi} = \bar{\mathbf{U}}_n\bar{\mathbf{T}}^{\dagger^H}\bar{\mathbf{T}}^\dagger\bar{\mathbf{U}}_n^H = \left[\begin{array}{cc}\underbrace{\bar{\mathbf{U}}_{n_1}\bar{\mathbf{T}}^{\dagger^H}\bar{\mathbf{T}}^\dagger\bar{\mathbf{U}}_{n_1}^H}_{\mathbf{\Phi}_{11}} & \underbrace{\mathbf{U}_{n_1}\bar{\mathbf{T}}^{\dagger^H}\bar{\mathbf{T}}^\dagger\bar{\mathbf{U}}_{n_2}^H}_{\mathbf{\Phi}_{12}}\\ \underbrace{\bar{\mathbf{U}}_{n_2}^H\bar{\mathbf{T}}^{\dagger^H}\bar{\mathbf{T}}^\dagger\bar{\mathbf{U}}_{n_1}^H}_{\mathbf{\Phi}_{12}^H} & \underbrace{\bar{\mathbf{U}}_{n_2}^H\bar{\mathbf{T}}^{\dagger^H}\bar{\mathbf{T}}^\dagger\bar{\mathbf{U}}_{n_2}^H}_{\mathbf{\Phi}_{22}}\end{array}\right]. \tag{2.91}$$

As $\bar{\mathbf{R}}\bar{\mathbf{U}}_n = \sigma^2\bar{\mathbf{U}}_n$, we have

$$\bar{\mathbf{R}}\mathbf{\Phi} = \mathbf{\Phi}\bar{\mathbf{R}} = \sigma^2\mathbf{\Phi}. \tag{2.92}$$

From (2.84) along with (2.92) it follows that

$$\mathrm{E}\left\{\delta\bar{\mathbf{R}}\mathbf{\Phi}\delta\bar{\mathbf{R}}\right\} = \mathrm{E}\left\{\hat{\bar{\mathbf{R}}}\mathbf{\Phi}\hat{\bar{\mathbf{R}}}\right\} - \sigma^4\mathbf{\Phi}. \tag{2.93}$$

Using (2.61) in $\mathrm{E}\left\{\hat{\bar{\mathbf{R}}}\boldsymbol{\Phi}\hat{\bar{\mathbf{R}}}\right\}$ yields

$$\begin{aligned}
\mathrm{E}\left\{\hat{\bar{\mathbf{R}}}\boldsymbol{\Phi}\hat{\bar{\mathbf{R}}}\right\} &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{E}\left\{\bar{\mathbf{x}}(i)\bar{\mathbf{x}}(i)^H\boldsymbol{\Phi}\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\right\} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\mathrm{E}\left\{\bar{\mathbf{x}}(i)\bar{\mathbf{x}}(i)^H\right\}\boldsymbol{\Phi}\mathrm{E}\left\{\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\right\} \\
&+ \frac{1}{N^2}\sum_{j=1}^{N}\mathrm{E}\left\{\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\boldsymbol{\Phi}\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\right\} \\
&= \frac{N-1}{N}\bar{\mathbf{R}}\boldsymbol{\Phi}\bar{\mathbf{R}} + \frac{1}{N}\mathrm{E}\left\{\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\boldsymbol{\Phi}\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\right\} \\
&= \frac{\sigma^4(N-1)}{N}\boldsymbol{\Phi} + \frac{1}{N}\mathrm{E}\left\{\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\boldsymbol{\Phi}\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\right\} \qquad (2.94)
\end{aligned}$$

where the second line follows from the fact that the received signals corresponding to different transmitted symbol intervals are independent from each other and the fourth line holds due to (2.92). Note that the columns of $\bar{\mathbf{U}}_n$ (and, hence, that of the matrix $\boldsymbol{\Phi}$) are orthogonal to the signal subspace, and we have that

$$\boldsymbol{\Phi}\bar{\mathbf{U}}_s = \boldsymbol{\Phi}\begin{bmatrix}\mathbf{W}\\ \mathbf{W}^*\end{bmatrix} = \mathbf{0}. \qquad (2.95)$$

Using (2.94) in (2.93), substituting the obtained expression into (2.90), and applying (2.95) to simplify the result, we have

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{1}{2N}\begin{bmatrix}\mathbf{w}_1\\ \mathbf{w}_1^*\end{bmatrix}^H\bar{\mathbf{U}}_s\bar{\boldsymbol{\Lambda}}_s^{-1}\bar{\mathbf{U}}_s^H\boldsymbol{\Upsilon}\bar{\mathbf{U}}_s\bar{\boldsymbol{\Lambda}}_s^{-1}\bar{\mathbf{U}}_s^H\begin{bmatrix}\mathbf{w}_1\\ \mathbf{w}_1^*\end{bmatrix} \qquad (2.96)$$

where

$$\boldsymbol{\Upsilon} = \mathrm{E}\left\{\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\boldsymbol{\Phi}\bar{\mathbf{x}}(j)\bar{\mathbf{x}}(j)^H\right\}. \qquad (2.97)$$

Using (2.49) to write $\bar{\mathbf{x}}(j)$ in terms of the transmitted symbols $\mathbf{b}(j)$ and noise $\mathbf{v}(j)$ and then taking into account (2.95) to simplify the resulting expression, after straightforward manipulations we obtain

$$\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}^{(1)} + \boldsymbol{\Upsilon}^{(2)} \qquad (2.98)$$

where

$$\boldsymbol{\Upsilon}^{(1)} = \mathrm{E}\left\{ \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \mathbf{b} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix}^H \boldsymbol{\Phi} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix} \mathbf{b}^H \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H \right\} \tag{2.99}$$

$$\boldsymbol{\Upsilon}^{(2)} = \mathrm{E}\left\{ \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix}^H \boldsymbol{\Phi} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix}^H \right\} \tag{2.100}$$

and the time index j has been omitted from $\mathbf{b}(j)$ and $\mathbf{v}(j)$ for the sake of brevity. We have

$$\begin{aligned}
\boldsymbol{\Upsilon}^{(1)} &= \mathrm{E}\left\{ \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix}^H \boldsymbol{\Phi} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix} \right\} \mathrm{E}\left\{ \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \mathbf{b}\mathbf{b}^H \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H \right\} \\
&= \mathrm{E}\left\{ \mathrm{tr}\left(\begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix}^H \boldsymbol{\Phi} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix} \right) \right\} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H \\
&= \mathrm{tr}\left(\mathrm{E}\left\{ \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}^* \end{bmatrix}^H \boldsymbol{\Phi} \right\} \right) \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H \\
&= \sigma^2 \mathrm{tr}(\boldsymbol{\Phi}) \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H \\
&= \sigma^2 \|\bar{\mathbf{T}}^\dagger\|_F^2 \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix}^H
\end{aligned} \tag{2.101}$$

where the last line directly follows from the definition of $\boldsymbol{\Phi}$ in (2.91). To obtain $\boldsymbol{\Upsilon}^{(2)}$, we can rewrite this matrix as

$$\boldsymbol{\Upsilon}^{(2)} = \begin{bmatrix} \boldsymbol{\Upsilon}_{11}^{(2)} & \boldsymbol{\Upsilon}_{12}^{(2)} \\ \boldsymbol{\Upsilon}_{12}^{(2)^H} & \boldsymbol{\Upsilon}_{22}^{(2)} \end{bmatrix} \tag{2.102}$$

where

$$\begin{aligned}
\boldsymbol{\Upsilon}_{11}^{(2)} &= \mathrm{E}\left\{\mathbf{v}\mathbf{v}^H \boldsymbol{\Phi}_{11} \mathbf{v}\mathbf{v}^H\right\} + \mathrm{E}\left\{\mathbf{v}\mathbf{v}^T \boldsymbol{\Phi}_{22} \mathbf{v}^* \mathbf{v}^H\right\} \\
&+ \mathrm{E}\left\{\mathbf{v}\mathbf{v}^T \boldsymbol{\Phi}_{12}^H \mathbf{v}\mathbf{v}^H\right\} + \mathrm{E}\left\{\mathbf{v}\mathbf{v}^H \boldsymbol{\Phi}_{12} \mathbf{v}^* \mathbf{v}^H\right\}
\end{aligned} \tag{2.103}$$

$$\begin{aligned}
\boldsymbol{\Upsilon}_{12}^{(2)} &= \mathrm{E}\left\{\mathbf{v}\mathbf{v}^H \boldsymbol{\Phi}_{12} \mathbf{v}^* \mathbf{v}^T\right\} + \mathrm{E}\left\{\mathbf{v}\mathbf{v}^H \boldsymbol{\Phi}_{11} \mathbf{v}\mathbf{v}^T\right\} \\
&+ \mathrm{E}\left\{\mathbf{v}\mathbf{v}^T \boldsymbol{\Phi}_{12}^H \mathbf{v}\mathbf{v}^T\right\} + \mathrm{E}\left\{\mathbf{v}\mathbf{v}^T \boldsymbol{\Phi}_{22} \mathbf{v}^* \mathbf{v}^T\right\}
\end{aligned} \tag{2.104}$$

$$\begin{aligned}\boldsymbol{\Upsilon}_{22}^{(2)} &= \mathrm{E}\left\{\mathbf{v}^*\mathbf{v}^T\boldsymbol{\Phi}_{22}\mathbf{v}^*\mathbf{v}^T\right\}+\mathrm{E}\left\{\mathbf{v}^*\mathbf{v}^H\boldsymbol{\Phi}_{11}\mathbf{v}\mathbf{v}^T\right\} \\ &+ \mathrm{E}\left\{\mathbf{v}^*\mathbf{v}^T\boldsymbol{\Phi}_{12}^H\mathbf{v}\mathbf{v}^T\right\}+\mathrm{E}\left\{\mathbf{v}^*\mathbf{v}^H\boldsymbol{\Phi}_{12}\mathbf{v}^*\mathbf{v}^T\right\}. \qquad (2.105)\end{aligned}$$

Noting to the fact that $\mathbf{v}$ is a white circular Gaussian random vector, for any arbitrary i, j, k and l, we have [30]

$$\begin{aligned}\mathrm{E}\{[\mathbf{v}]_i[\mathbf{v}]_j[\mathbf{v}]_k^*[\mathbf{v}]_l^*\} &= \mathrm{E}\{[\mathbf{v}]_i[\mathbf{v}]_k^*\}\mathrm{E}\{[\mathbf{v}]_j[\mathbf{v}]_l^*\}+\mathrm{E}\{[\mathbf{v}]_i[\mathbf{v}]_l^*\}\mathrm{E}\{[\mathbf{v}]_j[\mathbf{v}]_k^*\} \\ &= \sigma^2(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}) \qquad (2.106)\end{aligned}$$

and any other fourth-order moment is equal to zero. From the latter fact, it follows that the third and fourth terms in the right-hand side of (2.103) and (2.105) are equal to zero. The same is true for the second, third, and fourth terms in the right-hand side of (2.104). Using (2.106) to compute the remaining terms, it can be readily shown that

$$\boldsymbol{\Upsilon}_{11}^{(2)} = \sigma^4\left(\boldsymbol{\Phi}_{11}+\mathrm{tr}(\boldsymbol{\Phi}_{11})\mathbf{I}+\boldsymbol{\Phi}_{22}^T+\mathrm{tr}(\boldsymbol{\Phi}_{22})\mathbf{I}\right) \qquad (2.107)$$

$$\boldsymbol{\Upsilon}_{12}^{(2)} = \sigma^4\left(\boldsymbol{\Phi}_{12}+\boldsymbol{\Phi}_{12}^T\right) \qquad (2.108)$$

$$\boldsymbol{\Upsilon}_{22}^{(2)} = \sigma^4\left(\boldsymbol{\Phi}_{22}+\mathrm{tr}(\boldsymbol{\Phi}_{22})\mathbf{I}+\boldsymbol{\Phi}_{11}^T+\mathrm{tr}(\boldsymbol{\Phi}_{11})\mathbf{I}\right). \qquad (2.109)$$

Substituting (2.107)-(2.109) into (2.102) and using the fact that

$$\mathrm{tr}(\boldsymbol{\Phi}_{11})+\mathrm{tr}(\boldsymbol{\Phi}_{22})=\mathrm{tr}(\boldsymbol{\Phi})=\|\bar{\mathbf{T}}^\dagger\|_F^2 \qquad (2.110)$$

we obtain

$$\boldsymbol{\Upsilon}^{(2)}=\sigma^4\left(\boldsymbol{\Phi}+\bar{\boldsymbol{\Theta}}+\|\bar{\mathbf{T}}^\dagger\|_F^2\mathbf{I}\right) \qquad (2.111)$$

where

$$\bar{\boldsymbol{\Theta}}=\begin{bmatrix}\boldsymbol{\Phi}_{22}^T & \boldsymbol{\Phi}_{12}^T \\ \boldsymbol{\Phi}_{12}^* & \boldsymbol{\Phi}_{11}^T\end{bmatrix}. \qquad (2.112)$$

From (2.101) and (2.111) along with (2.50), it follows that

$$\boldsymbol{\Upsilon}=\sigma^4\boldsymbol{\Phi}+\sigma^4\bar{\boldsymbol{\Theta}}+\sigma^2\|\bar{\mathbf{T}}^\dagger\|_F^2\bar{\mathbf{R}}. \qquad (2.113)$$

To proceed with the proof, first we need to show that the columns of $\bar{\boldsymbol{\Theta}}$ are orthogonal to the signal subspace. Using (2.91) and (2.112), the matrix $\bar{\boldsymbol{\Theta}}$ can be written as

$$\bar{\boldsymbol{\Theta}}=\begin{bmatrix}\bar{\mathbf{U}}_{n_2}^* \\ \bar{\mathbf{U}}_{n_1}^*\end{bmatrix}\bar{\mathbf{T}}^{\dagger^T}\bar{\mathbf{T}}^*\begin{bmatrix}\bar{\mathbf{U}}_{n_2}^* \\ \bar{\mathbf{U}}_{n_1}^*\end{bmatrix}^H. \qquad (2.114)$$

From (2.55) we have

$$\bar{\mathbf{U}}_{n_1}^H \mathbf{W} + \bar{\mathbf{U}}_{n_2}^H \mathbf{W}^* = \mathbf{0}. \tag{2.115}$$

Applying the complex conjugate to both sides of (2.115), we have

$$\begin{bmatrix} \bar{\mathbf{U}}_{n_2}^* \\ \bar{\mathbf{U}}_{n_1}^* \end{bmatrix}^H \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} = \mathbf{0}. \tag{2.116}$$

From (2.116), it follows that the columns of $\begin{bmatrix} \bar{\mathbf{U}}_{n_2}^* \\ \bar{\mathbf{U}}_{n_1}^* \end{bmatrix}$ are in $\mathtt{range}(\bar{\mathbf{U}}_n)$, or, equivalently, there exists a matrix $\mathbf{P}$ such that

$$\begin{bmatrix} \bar{\mathbf{U}}_{n_2}^* \\ \bar{\mathbf{U}}_{n_1}^* \end{bmatrix} = \bar{\mathbf{U}}_n \mathbf{P}. \tag{2.117}$$

Using (2.116) and (2.117) along with (2.114), we obtain

$$\bar{\mathbf{\Theta}} \bar{\mathbf{U}}_s = \bar{\mathbf{\Theta}} \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} = \mathbf{0}. \tag{2.118}$$

Substituting (2.113) into (2.96) and taking (2.95) and (2.118) to account, we have

$$\mathrm{E}\left\{\|\delta \mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2 \|\bar{\mathbf{T}}^\dagger\|_F^2}{2N} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_1^* \end{bmatrix}^H \bar{\mathbf{U}}_s \bar{\mathbf{\Lambda}}_s^{-1} \bar{\mathbf{U}}_s^H \bar{\mathbf{R}} \bar{\mathbf{U}}_s \bar{\mathbf{\Lambda}}_s^{-1} \bar{\mathbf{U}}_s^H \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_1^* \end{bmatrix}. \tag{2.119}$$

Equation (2.80) directly follows from (2.119) and from the fact that

$$\bar{\mathbf{U}}_s^H \bar{\mathbf{R}} \bar{\mathbf{U}}_s = \bar{\mathbf{\Omega}}_s. \tag{2.120}$$

To prove (2.83), note that if (2.23) holds, then we have

$$\begin{bmatrix} \mathbf{w}_i \\ \mathbf{w}_i^* \end{bmatrix}^H \begin{bmatrix} \mathbf{w}_j \\ \mathbf{w}_j^* \end{bmatrix} = 2\|\mathbf{w}_i\|^2 \delta_{ij} \tag{2.121}$$

and, hence, the normalized version of $\begin{bmatrix} \mathbf{w}_i \\ \mathbf{w}_i^* \end{bmatrix}$ can be considered as the i-th column of

$$\bar{\mathbf{U}}_s = \begin{bmatrix} \frac{\mathbf{w}_1}{\sqrt{2}\|\mathbf{w}_1\|} & \frac{\mathbf{w}_2}{\sqrt{2}\|\mathbf{w}_2\|} & \cdots & \frac{\mathbf{w}_K}{\sqrt{2}\|\mathbf{w}_K\|} \\ \frac{\mathbf{w}_1^*}{\sqrt{2}\|\mathbf{w}_1\|} & \frac{\mathbf{w}_2^*}{\sqrt{2}\|\mathbf{w}_2\|} & \cdots & \frac{\mathbf{w}_K^*}{\sqrt{2}\|\mathbf{w}_K\|} \end{bmatrix}. \tag{2.122}$$

In such a scenario, we also have

$$\bar{\mathbf{\Lambda}}_s = \mathrm{diag}\left(2A_1^2\|\mathbf{w}_1\|^2, 2A_2^2\|\mathbf{w}_2\|^2, \ldots, 2A_K^2\|\mathbf{w}_K\|^2\right). \tag{2.123}$$

Substituting (2.122) and (2.123) into the right-hand side of (2.119) and using (2.121) to simplify the resulting expression, we obtain

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2\|\bar{\mathbf{T}}^\dagger\|_F^2}{2A_1^2N}\left(1 + \frac{\sigma^2}{2A_1^2\|\mathbf{w}_1\|^2}\right). \tag{2.124}$$

If (2.27) holds, (2.83) directly follows from (2.124). This completes our proof. □

Equation (2.83) explicitly expresses the MSE of the estimated channel vector in terms of *environmental parameters* such as the noise power, the power of the user-of-interest, and the number of data samples. The effect of the matrix $\bar{\mathbf{T}}$ on the MSE of the estimated channel vector can be explained as follows. Since $\mathbf{h}_1$ is a solution to (2.71), it follows that

$$\bar{\mathbf{T}}\begin{bmatrix}\mathbf{h}_1\\ \mathbf{h}_1^*\end{bmatrix} = \begin{bmatrix}\bar{\mathbf{U}}_{n_1}^H\boldsymbol{\mathcal{C}}_1 & \bar{\mathbf{U}}_{n_2}^H\boldsymbol{\mathcal{C}}_1^*\end{bmatrix}\begin{bmatrix}\mathbf{h}_1\\ \mathbf{h}_1^*\end{bmatrix} = \mathbf{0} \tag{2.125}$$

and, hence, the $[2(L_c - L + 1) - K] \times 2L$ matrix $\bar{\mathbf{T}}$ is low rank. If

$$\mathtt{rank}(\bar{\mathbf{T}}) = 2L - 1 \tag{2.126}$$

then, according to (2.125), the null-space of $\bar{\mathbf{T}}$ is spanned by $[\mathbf{h}_1^T \ \mathbf{h}_1^H]^T$, and, up to an arbitrary scaling factor, $\mathbf{h}_1$ is the unique solution to (2.71). In such a case, (2.83) can be written as

$$\mathrm{E}\{\|\delta\mathbf{h}_1\|^2\} \approx \frac{\sigma^2\sum_{i=1}^{2L-1}\varrho_i{}^{-2}}{2A_1^2N} \tag{2.127}$$

where $\varrho_1 \geq \varrho_2 \geq \ldots \geq \varrho_{2L-1} > 0$ are the positive singular values of $\bar{\mathbf{T}}$. Equation (2.127) clearly shows that even if $\mathbf{h}_1$ is uniquely identifiable, the channel estimation performance may substantially degrade if the smallest positive eigenvalue ϱ_{2L-1} is close to zero. Similar observation can be made for the LX algorithm [47].

2.4.3 Extension to temporally oversampled received signals

If temporal oversampling is feasible, the proposed algorithm can be applied to the oversampled received data. This results in further increase of the maximal admissible values of K and L as well as improvements in the estimation performance. Denoting

$$\bar{\mathbf{x}}_\kappa(n) = [\mathbf{x}_\kappa^T(n)\ \mathbf{x}_\kappa^H(n)]^T \tag{2.128}$$

we have that

$$\bar{\mathbf{R}}_\kappa = \mathrm{E}\{\bar{\mathbf{x}}_\kappa(n)\bar{\mathbf{x}}_\kappa^H(n)\} = \begin{bmatrix} \mathbf{W}_\kappa \\ \mathbf{W}_\kappa^* \end{bmatrix} \begin{bmatrix} \mathbf{W}_\kappa \\ \mathbf{W}_\kappa^* \end{bmatrix}^H + \sigma^2\mathbf{I}. \tag{2.129}$$

The matrix $\bar{\mathbf{R}}_\kappa$ can be eigendecomposed as

$$\bar{\mathbf{R}}_\kappa = \begin{bmatrix} \bar{\mathbf{U}}_{s\kappa} & \bar{\mathbf{U}}_{n\kappa} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{\Omega}}_{s\kappa} & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{U}}_{s\kappa}^H \\ \bar{\mathbf{U}}_{n\kappa}^H \end{bmatrix} \tag{2.130}$$

where the $K \times K$ matrix $\bar{\mathbf{\Omega}}_{s\kappa}$ contains the K largest eigenvalues of $\bar{\mathbf{R}}_\kappa$ and the $2\kappa(L_c - L + 1) \times K$ matrix $\bar{\mathbf{U}}_s$ contains the corresponding orthonormal eigenvectors. Let us denote the matrices containing the first $\kappa(L_c - L + 1)$ rows and the remaining last $\kappa(L_c - L + 1)$ rows of $\bar{\mathbf{U}}_{n\kappa}$ as $\bar{\mathbf{U}}_{n\kappa_1}$ and $\bar{\mathbf{U}}_{n\kappa_2}$, respectively. Using similar procedure as in Section 2.4, it can be readily shown that

$$\begin{bmatrix} \mathrm{Re}(\bar{\mathbf{T}}_{\kappa_1}) + \mathrm{Re}(\bar{\mathbf{T}}_{\kappa_2}) & \mathrm{Im}(\bar{\mathbf{T}}_{\kappa_2}) - \mathrm{Im}(\bar{\mathbf{T}}_{\kappa_1}) \\ \mathrm{Im}(\bar{\mathbf{T}}_{\kappa_1}) + \mathrm{Im}(\bar{\mathbf{T}}_{\kappa_2}) & \mathrm{Re}(\bar{\mathbf{T}}_{\kappa_1}) - \mathrm{Re}(\bar{\mathbf{T}}_{\kappa_2}) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\mathbf{h}_{1\kappa}) \\ \mathrm{Im}(\mathbf{h}_{1\kappa}) \end{bmatrix} = \mathbf{0} \tag{2.131}$$

where

$$\begin{aligned} \bar{\mathbf{T}}_{\kappa_1} &= \bar{\mathbf{U}}_{n\kappa_1}^H \mathcal{C}_{1\kappa} \\ \bar{\mathbf{T}}_{\kappa_2} &= \bar{\mathbf{U}}_{n\kappa_2}^H \mathcal{C}_{1\kappa}^*. \end{aligned} \tag{2.132}$$

The linear system in (2.131) consists of $4\kappa(L_c - L + 1) - 2K$ real equations and $2\kappa L$ real unknown variables. Hence, the unique nontrivial solution to (2.131) requires that

$$3\kappa L + K \leq 2\kappa(L_c + 1). \tag{2.133}$$

Comparing (2.133) with (2.46), it can be observed that the restriction on the maximum number of users and the channel length is further relaxed. Hence, the improvements of oversampling come on top of those achieved by means of the proposed approach.

2.5 Simulations

To compare the performances of the proposed algorithm and the LX technique and to validate our analytical results, computer simulations have been carried out. In all numerical examples, $L_c = 40$ and the spreading sequence associated with each user has been randomly drawn from the binary set of $\{-1, +1\}$ and then fixed throughout all examples. Similarly, the entries of the channel vectors have been randomly and independently drawn from a zero-mean complex Gaussian process and then have been normalized so that $\|\mathbf{h}_k\| = 1$ $(k = 1, \ldots, K)$ and fixed throughout all examples. In each simulation run, the entries of the noise vector and the transmitted symbols have been randomly and independently drawn from a zero-mean white circular multivariate Gaussian process and the BPSK constellation, respectively. Throughout the simulations, we assume that all users have identical powers. All simulation curves are averaged using 500 simulation runs.

Fig. 2.1 shows the experimental MSEs of both algorithms as well as theoretical MSEs (2.80) and (2.83) versus the number of users. In this figure, the channel length $L = 8$ is selected and $N = 80$ data vectors have been used to estimate the covariance matrices $\mathbf{R}$ and $\bar{\mathbf{R}}$ for the LX algorithm and the proposed algorithm, respectively. The SNR of all users is assumed to be equal to 10 dB. Substituting the selected values of L_c and L in (2.18), it can be concluded that the LX algorithm cannot properly operate if the number of users is larger than 25. Moreover, if $K \geq L_c - L + 1 = 33$, then the LX algorithm becomes completely overloaded (as the dimension of the noise subspace shrinks to zero). These theoretical observations are fully validated by Fig. 2.1. At the same time, as can be observed from this figure, the proposed technique is able to identify the desired channel vector in a system with up to 50 active users, i.e., it can properly operate in more heavily loaded environments than the LX algorithm. Furthermore, as can be seen from Fig. 2.1,

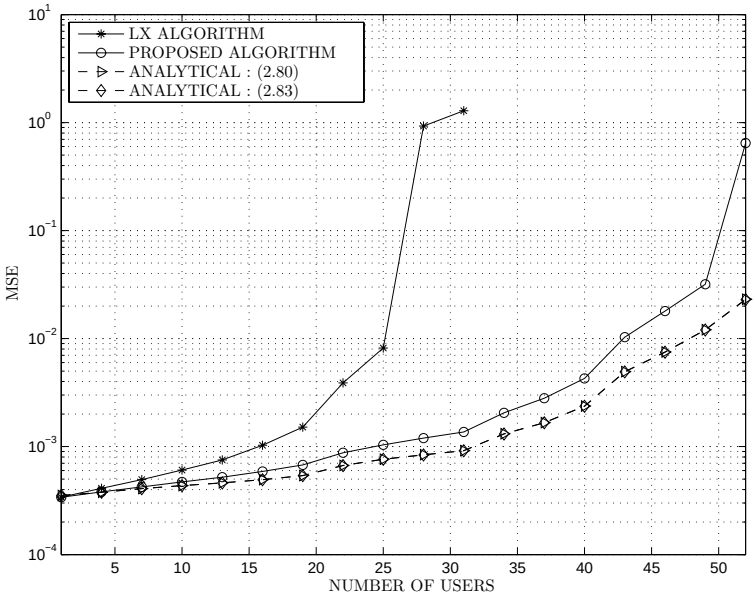

Figure 2.1: Channel estimation MSEs versus the number of users.

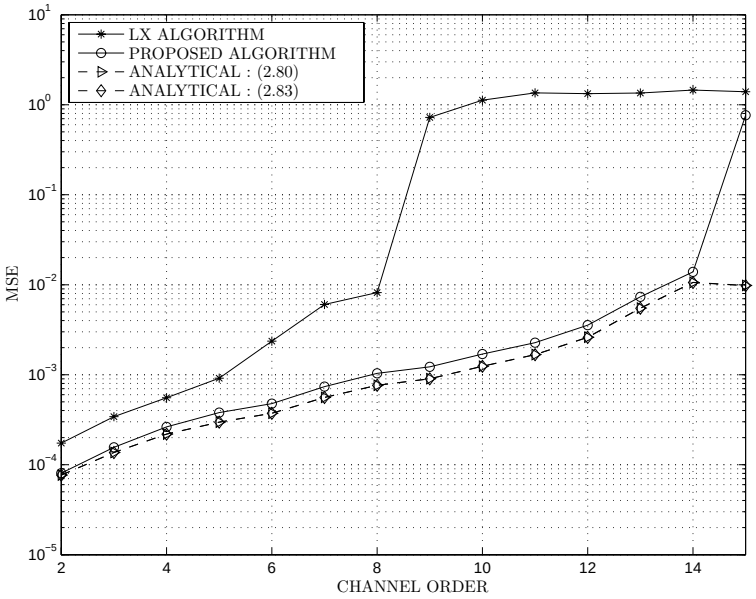

Figure 2.2: Channel estimation MSEs versus the channel order.

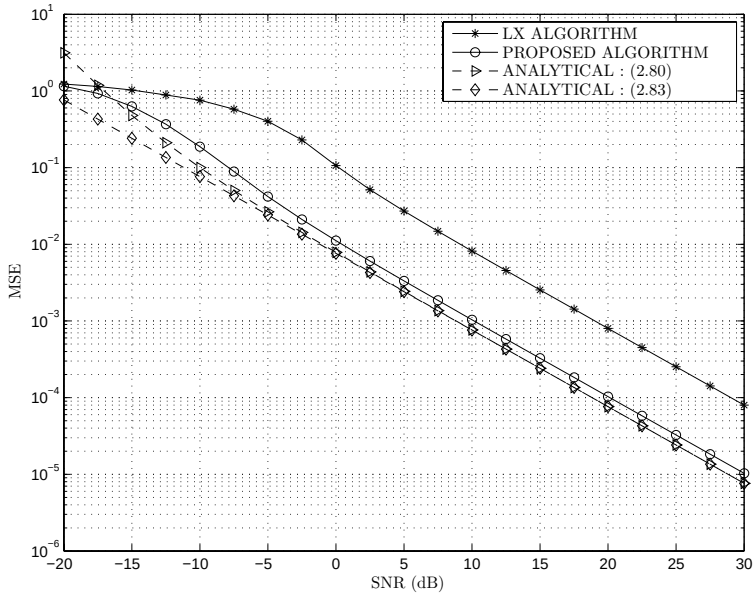

Figure 2.3: Channel estimation MSEs versus the SNR.

even within the operating range of the LX algorithm ($K \leq 25$), the latter technique is outperformed by the proposed algorithm. Fig. 2.1 also clearly shows that if the number of users is not very large, the analytical MSE curves obtained from (2.80) and (2.83) follow the experimental MSE curve of the proposed algorithm with a good accuracy. As the MSE expressions (2.80) and (2.83) are derived using the first-order perturbation theory, their validity requires a small perturbation assumption. Therefore, if the number of active users (and, consequently, the MSE of the channel vector estimate) become large, then, as confirmed by Fig. 2.1, equations (2.80) and (2.83) cannot accurately predict the MSE values. Moreover, it can be observed from Fig. 2.1 that the analytical MSE curves of (2.80) and (2.83) are almost indistinguishable from each other. This is due to the fact that the SNR of the user-of-interest is high enough, and, hence, the assumption (2.27) holds.

Fig. 2.2 shows the experimental MSEs of both algorithms tested along with the theoretical MSEs (2.80) and (2.83) versus the channel length L. In this figure, the user SNRs and the observation length N are identical to their values in Fig. 2.1, and the number of users $K = 25$ is fixed. From (2.18), we can conclude that using the LX algorithm, it is not

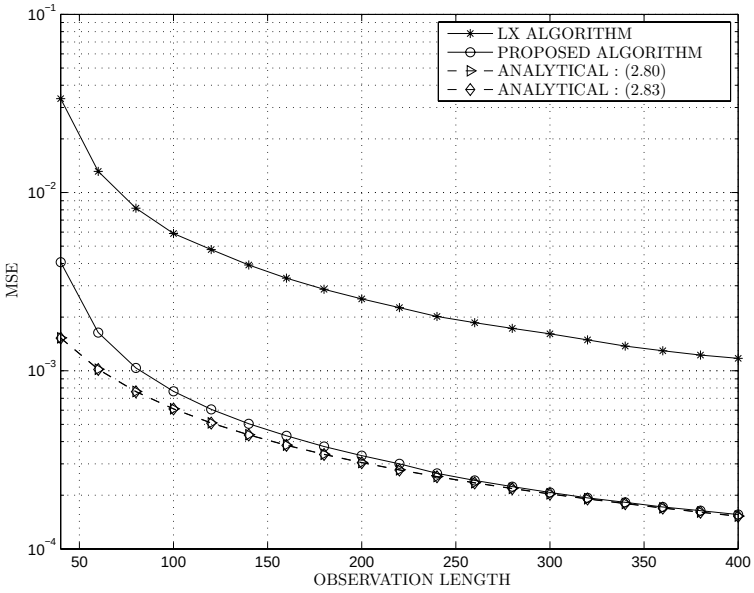

Figure 2.4: Channel estimation MSEs versus the observation length.

possible to identify the channel with the length $L > 8$. This fact is fully verified by Fig. 2.2, from which we also see that the proposed algorithm reliably estimates the channel vector with the length up to $L = 14$. Moreover, our algorithm outperforms the LX algorithm over the whole range of L tested. Fig. 2.2 also further verifies the validity of the MSE expressions (2.80) and (2.83) that follow the experimental MSE curve of the proposed algorithm with a high accuracy.

Fig. 2.3 shows the experimental MSEs of both algorithms along with the analytical MSE curves obtained from (2.80) and (2.83) versus the user SNR for $N = 80$, $L = 8$, and $K = 25$. Similar to the previous two figures, substantial performance improvements of the proposed algorithm as compared to the LX algorithm can be observed. The accuracy of the MSE expressions (2.80) and (2.83) over a wide range of sufficiently high SNRs is also confirmed.

Fig. 2.4 displays the experimental and analytical MSEs versus the observation length N for SNR = 10 dB, $K = 25$, and $L = 8$. It can be seen from this figure that, although both the experimental MSE curves converge with similar rates, the proposed algorithm requires smaller sample size as compared to the LX algorithm. The figure also shows that, as may

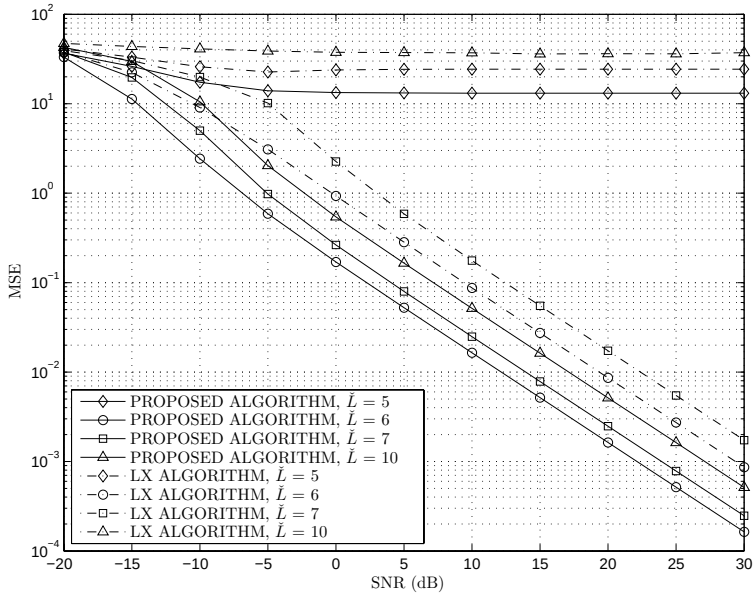

Figure 2.5: Signature estimation MSEs versus the SNR for $\check{L} = 5, 6, 7, 10$.

be expected, the analytical MSE curves get closer to the experimental MSE curve of our algorithm when increasing the observation length.

The effects of the channel order mismatch on both algorithms is examined in Fig. 2.5 for $N = 80$ and $K = 25$. The actual channel length is $L = 6$ while four different values of $\check{L}$ ($\check{L} = 5; 6; 7; 10$) are used as the presumed channel length in both algorithms. Since for $\check{L} = 5, 7$, and 10 the length of the channel vector estimate is different from that of the actual channel vector, the estimation error vector $\delta\mathbf{h}_1 = \hat{\mathbf{h}}_1 - \mathbf{h}_1$ is not properly defined and one should resort to some other criterion to evaluate the estimation performance. In this numerical example, for each value of $\check{L}$ the channel vector estimate $\hat{\mathbf{h}}_1$ is used in (1.12) to obtain the signature estimate $\underline{\hat{\mathbf{w}}}_1 = \underline{\boldsymbol{\mathcal{C}}}_1\hat{\mathbf{h}}_1$. Fig. 2.5 shows the signature MSE

$$\mathrm{E}\{\|\delta\mathbf{w}_1\|^2\} = \mathrm{E}\{\|\underline{\hat{\mathbf{w}}}_1 - \underline{\mathbf{w}}_1\|^2\} \tag{2.134}$$

versus SNR for both algorithms and all the presumed values of $\check{L}$. As can be observed from the signature MSE curves corresponding to $\check{L} = 5$, both algorithms are very sensitive to underestimation of the channel length. The figure also shows that if $\check{L} = L = 6$, both algorithms provide their best estimation performance, and, at the same time, our

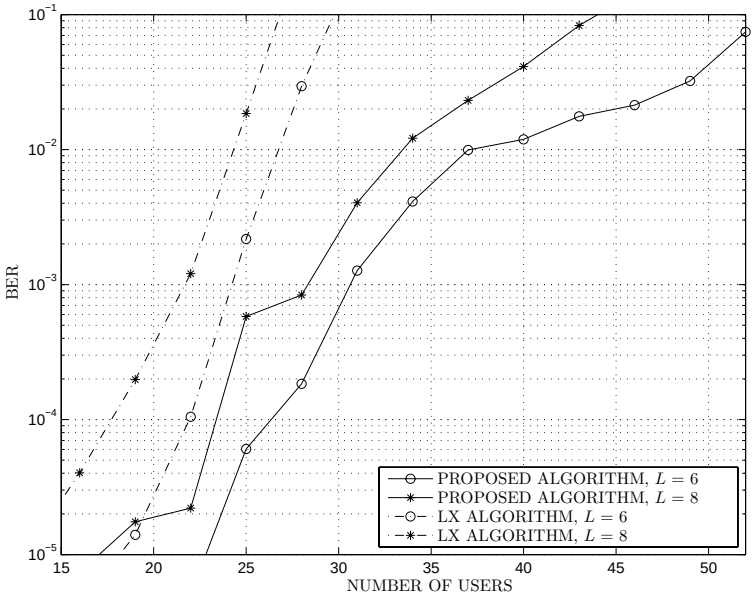

Figure 2.6: BERs of the MMSE receivers versus the number of users.

algorithm substantially outperforms the LX algorithm. The MSE curves corresponding to $\check{L} = 7$ suggest that both algorithms are quite robust to overestimation of the channel length. However, as it can be observed from the figure at $\check{L} = 10$, our algorithm still can provide reliable estimates while the LX algorithm completely breaks down. It is due to the fact that according to (2.18), $\check{L} = 10$ exceeds the maximal channel length admissible for the LX algorithm. Interestingly, the proposed algorithm with erroneous $\check{L} = 10$ even outperforms the LX algorithm with the perfect knowledge of the channel length ($\check{L} = L = 6$). This observation further validates advantages of the proposed algorithm in practical scenarios where often the exact knowledge of the channel vector length is not available or is prohibitively difficult to obtain.

Fig. 2.6 compares the BER performance of the MMSE receiver which uses the signature vector estimate obtained from the proposed technique against that of the MMSE receiver which uses the signature vector estimate of the LX algorithm. The true data covariance matrix $\mathbf{R}$ is used in both MMSE receivers to make the resulting BER curves only representing the effect of the signature vector estimation techniques on the symbol detection

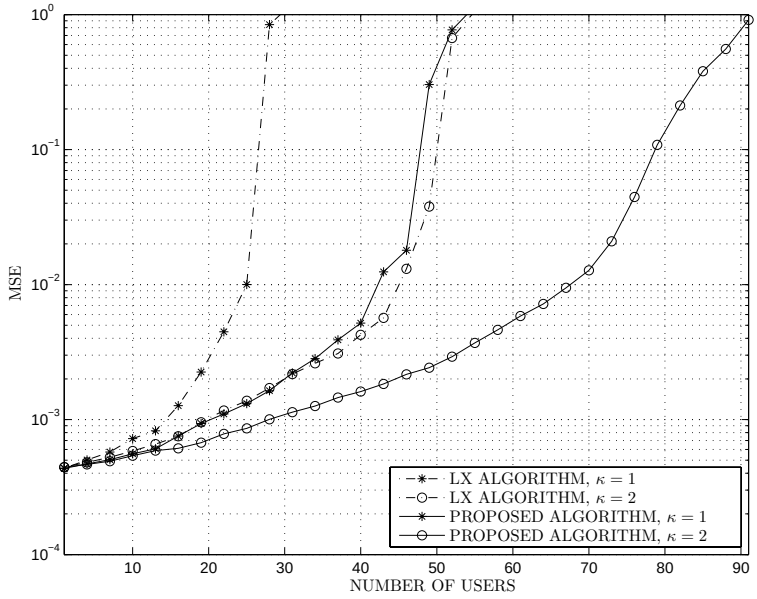

Figure 2.7: Channel estimation MSEs versus the number of users in the the presence of oversampling.

performance. The BER curves are plotted versus the number of users for SNR = 0 dB and two different channel lengths of $L = 6$ and $L = 8$. This figure clearly demonstrates the advantage of using the proposed estimation technique in conjunction with the MMSE multiuser receiver.

Figs. 2.7 and 2.8 illustrate the performance of the proposed and LX algorithms in the case of oversampling for the same scenarios as examined in Figs. 2.1 and 2.2, respectively. The oversampling factor κ in Figs. 2.7 and 2.8 is varied from one (no oversampling) to two (doubling the sampling rate). These figures show that, as it may be expected, the performances of both algorithms can be improved by oversampling. Our algorithm can further improve the performance with respect to the oversampled LX algorithm because the improvements achieved by our approach come on top of the improvements provided by oversampling.

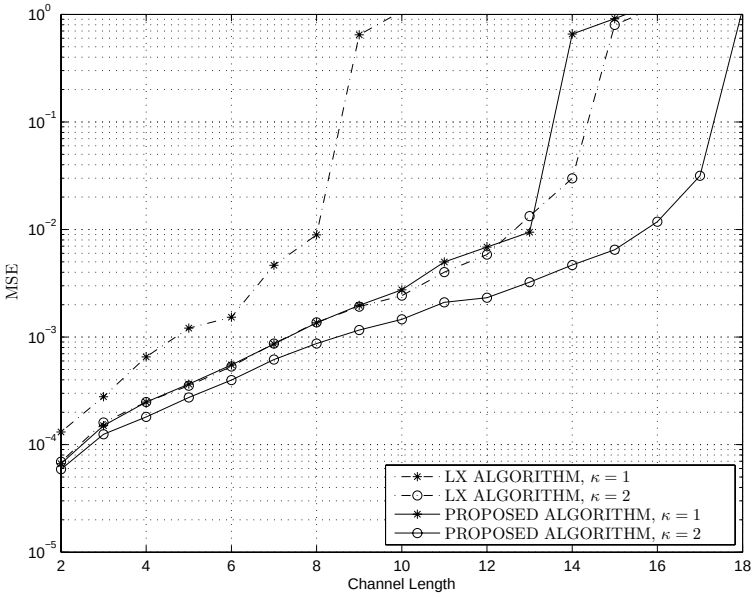

Figure 2.8: Channel estimation MSEs versus the channel order in the presence of oversampling.

2.6 Chapter summary

A new blind subspace-based algorithm for signature waveform estimation in DS-CDMA systems has been proposed. Although our algorithm uses the key idea of the popular LX technique, it additionally exploits non-circularity of the BPSK transmitted symbols and circularity of noise to increase the dimension of the observation space twice with respect to the LX algorithm while keeping the dimension of the signal subspace unchanged. As a result, the proposed technique has a substantially improved performance as compared to the LX method. Moreover, the proposed algorithm has been shown to be applicable to scenarios with an increased number of active users which are allowed to have longer channels as compared to the scenarios that can be properly treated by the LX technique. These results have validated our theoretical identifiability analysis of the proposed technique which has shown that it enjoys more relaxed identifiability conditions than the LX algorithm. Analytical expressions of the MSE of the proposed estimation algorithm have been also derived and validated by numerical examples.

Chapter 3

Signature estimation in the presence of correlated noise

3.1 Introduction

A typical assumption used in the blind subspace-based signature waveform estimation techniques is that the additive ambient noise is temporally white, and, hence, the signal subspace can be extracted using eigendecomposition of the data covariance matrix. However, in practice this assumption may be violated [8], [90]. It is well-known that in the presence of correlated noise, the signal subspace can not be identified from the subspace spanned by the eigenvectors associated with the largest eigenvalues of the data covariance matrix. Therefore, some alternative approaches should be employed to identify the signal subspace in the correlated noise case.

One of such approaches has been proposed by Wang and Poor [90]. Their technique is based on the assumption that the receiver contains two well-separated antennas such that the receiver noise is spatially white. Using this fact, the signal subspace can be obtained from the cross-correlation between the data received by different antennas. Hereafter, we refer to this technique as the WP (Wang and Poor) algorithm.

As deploying two widely separated antennas in the current mobile transceivers may be practically infeasible, another technique that addresses the aforementioned problem has been proposed by Buzzi and Poor [11] to identify the signal subspace using just a single

antenna at the receiver. This technique is based on the assumption that the noise is a circular Gaussian process while the transmitted symbols are non-circular (BPSK) signals. In such a case, it has been shown in [11] that the signal subspace can be directly identified using the singular value decomposition of the data pseudo-covariance matrix. Hereafter, we refer to the technique as the BP (Buzzi and Poor) algorithm.

Although the performance of the conventional (white noise assumption based) signature waveform estimation techniques has been frequently investigated in the literature [47], [76], [99], [106], only a little effort has been made to analyze the performance of the estimation algorithms proposed for the unknown correlated noise scenario. In this chapter (see also [114], [115]), we use the first-order perturbation theory to derive approximate expressions for the MSE of the channel vector estimates obtained by the WP and BP algorithms. Under several mild assumptions, simple high SNR approximations of these MSE expressions are also obtained. The derived MSE expressions clarify how the performance of the algorithms depends on the physical parameters such as the number of data samples, the received power of the user-of-interest, and the noise covariance matrix. The effects of the spreading factor and the channel length on the performance of the WP and BP algorithms are also studied. It is shown that the performance of these algorithms depends not only on SNR but also on the direction of the eigenvectors of the noise covariance matrix. To clarify this fact, we fix the eigenvalues of the noise covariance matrix and find the sets of eigenvectors which maximize (minimize) the MSE of the channel vector estimates. Moreover, over all noise covariance matrices with fixed trace, we obtain those which correspond to the extremal values of the MSE. It is shown that both the maximum and the minimum values of the MSE are obtained when the noise covariance matrix is rank deficient. As the trace of the noise covariance matrix is equal to the average noise power, the latter observation shows that the performance of the algorithms is more sensitive to a low-rank interference than to a full-rank noise with the same average power. We also show that in the presence of white noise, the performances of the WP and BP algorithms are identical to that of the conventional LX algorithm [47] that was developed for the white noise case.

Assuming that the SNR is high and the WP algorithm is used to estimate the channel

vector between the user of interest and the first antenna, we prove that the estimation performance is independent from the noise covariance matrix and the user received power at the second antenna. We use the latter property to show that when the receiver is equipped with multiple antennas, the second antenna can be arbitrarily chosen at high SNRs.

The rest of this chapter is organized as follows. A brief overview of the WP and BP algorithms is provided in Section 3.2. Section 3.3 presents our main theoretical results on the performance of the WP and BP algorithms. Simulation results validating our analysis are presented in Section 3.4. A summary of the chapter is given in Section 3.5.

3.2 Blind channel estimation

3.2.1 WP algorithm

The WP algorithm assumes that the receiver is equipped with two well-separated antennas such that the noise is *spatially uncorrelated* between them. Similar to (2.6), the ISI-free part of the sampled received data vectors at each antenna are given by

$$\mathbf{x}^{(i)}(n) = \mathbf{W}^{(i)}\mathbf{b}(n) + \mathbf{v}^{(i)}(n), \qquad i = 1, 2 \tag{3.1}$$

where i is the antenna index,

$$\mathbf{W}^{(i)} = \left[A_1^{(i)}\mathbf{w}_1^{(i)}, A_2^{(i)}\mathbf{w}_2^{(i)}, \ldots, A_K^{(i)}\mathbf{w}_K^{(i)}\right] \tag{3.2}$$

$\mathbf{v}^{(i)}(n)$ is noise at the ith antenna, and $A_k^{(i)}$ and

$$\mathbf{w}_k^{(i)} = \boldsymbol{\mathcal{C}}_k\mathbf{h}_k^{(i)} \tag{3.3}$$

are the amplitude and the signature vector of the kth user at the ith antenna, respectively. The covariance matrix corresponding to the sampled received data vector at each antenna is given by [90]

$$\mathbf{R}^{(i)} = \mathrm{E}\left\{\mathbf{x}^{(i)}(n){\mathbf{x}^{(i)}}^H(n)\right\} = \mathbf{W}^{(i)}{\mathbf{W}^{(i)}}^H + \boldsymbol{\Sigma}_v^{(i)} \qquad i = 1, 2 \tag{3.4}$$

where

$$\boldsymbol{\Sigma}_v^{(i)} = \mathrm{E}\left\{\mathbf{v}^{(i)}(n){\mathbf{v}^{(i)}}^H(n)\right\}. \tag{3.5}$$

As noise is uncorrelated between the antennas, we have [90]

$$\begin{aligned}\mathbf{R}^{(12)} &= \mathrm{E}\left\{\mathbf{x}^{(1)}(n)\mathbf{x}^{(2)\,H}(n)\right\} = \mathbf{W}^{(1)}\mathbf{W}^{(2)\,H} \\ &= \begin{bmatrix} \mathbf{U}_s^{(1)} & \mathbf{U}_n^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{\Omega}_s^{(12)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^{(2)\,H} \\ \mathbf{U}_n^{(2)\,H} \end{bmatrix}\end{aligned} \tag{3.6}$$

where the right-hand side of (3.6) is SVD of $\mathbf{R}^{(12)}$. It is clear that

$$\mathtt{range}(\mathbf{U}_s^{(1)}) \;=\; \mathtt{range}(\mathbf{W}^{(1)}) \tag{3.7}$$

$$\mathtt{range}(\mathbf{U}_s^{(2)}) \;=\; \mathtt{range}(\mathbf{W}^{(2)}). \tag{3.8}$$

For the sake of simplicity but without any loss of generality, assume that we are interested in the channel vector between the first user and the first antenna. Then, we have [90]

$$\mathbf{U}_n^{(1)\,H}\mathbf{w}_1^{(1)} = \mathbf{T}_1^{(1)}\mathbf{h}_1^{(1)} = \mathbf{0} \tag{3.9}$$

where

$$\mathbf{T}_1^{(1)} = \mathbf{U}_n^{(1)\,H}\boldsymbol{\mathcal{C}}_1 \tag{3.10}$$

is an $L_c - L + 1 - K \times L$ matrix. If

$$\mathtt{rank}(\mathbf{T}_1^{(1)}) = L - 1 \tag{3.11}$$

then, up to an arbitrary phase rotation, $\mathbf{h}_1^{(1)}$ is the unique non-trivial solution to (3.9) subject to $\|\mathbf{h}_1^{(1)}\| = 1$ [90]. In practice, $\mathbf{R}^{(12)}$ can be estimated as

$$\hat{\mathbf{R}}^{(12)} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}^{(1)}(n)\mathbf{x}^{(2)\,H}(n) \tag{3.12}$$

which results in the following estimate of $\mathbf{h}_1^{(1)}$ [90]:

$$\hat{\mathbf{h}}_1^{(1)} = \mho\left(\boldsymbol{\mathcal{C}}_1^H\hat{\mathbf{U}}_n^{(1)}\hat{\mathbf{U}}_n^{(1)\,H}\boldsymbol{\mathcal{C}}_1\right) \tag{3.13}$$

where $\hat{\mathbf{U}}_n^{(1)}$ consists of the left singular vectors associated with the $L_c - L + 1 - K$ smallest singular values of $\hat{\mathbf{R}}^{(12)}$.

3.2.2 BP algorithm

Another approach to solve the problem of channel estimation in the presence of unknown correlated noise has been proposed in [11]. Without requiring the second antenna, this algorithm is based on the assumption that the transmitted symbols are drawn from the BPSK constellation and the noise is a circular Gaussian process. Let

$$\tilde{\mathbf{R}} = \mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}^T(n)\right\} \tag{3.14}$$

be the pseudo-covariance matrix of the ISI-free part of the sampled received data. Using (2.6) along with (2.47), we have [11]

$$\tilde{\mathbf{R}} = \mathbf{W}\mathbf{W}^T = \begin{bmatrix} \tilde{\mathbf{U}}_s & \tilde{\mathbf{U}}_n \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Omega}}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{V}}_s^H \\ \tilde{\mathbf{V}}_n^H \end{bmatrix} \tag{3.15}$$

where $\tilde{\mathbf{\Omega}}_s$ is a diagonal matrix whose diagonal elements are the non-zero singular values of $\tilde{\mathbf{R}}$ and the columns of $\tilde{\mathbf{U}}_s$ are the corresponding singular vectors. It is easy to show that [11]

$$\mathtt{range}(\tilde{\mathbf{U}}_s) = \mathtt{range}(\mathbf{W}) \tag{3.16}$$

and, hence,

$$\tilde{\mathbf{U}}_n^H \mathbf{w}_1 = \tilde{\mathbf{T}}_1 \mathbf{h}_1 = \mathbf{0} \tag{3.17}$$

where

$$\tilde{\mathbf{T}}_1 = \tilde{\mathbf{U}}_n^H \boldsymbol{\mathcal{C}}_1. \tag{3.18}$$

It can be observed that $\tilde{\mathbf{T}}_1$ is an $(L_c - L + 1 - K) \times L$ matrix and the unique identification of $\mathbf{h}_1$ from (3.17) requires that [11]

$$\mathtt{rank}(\tilde{\mathbf{T}}_1) = L - 1. \tag{3.19}$$

In practice, similar to the LX and WP algorithms, $\mathbf{h}_1$ can be estimated by

$$\hat{\mathbf{h}}_1 = \mho\left(\boldsymbol{\mathcal{C}}_1^H \hat{\tilde{\mathbf{U}}}_n \hat{\tilde{\mathbf{U}}}_n^H \boldsymbol{\mathcal{C}}_1\right) \tag{3.20}$$

where $\hat{\tilde{\mathbf{U}}}_n$ is the matrix containing the left singular vectors associated with the $L_c-L+1-K$ least singular values of $\hat{\tilde{\mathbf{R}}}$, and

$$\hat{\tilde{\mathbf{R}}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}(n)^T \tag{3.21}$$

is the sample estimate of $\tilde{\mathbf{R}}$.

3.3 Performance analysis

3.3.1 Performance analysis of the WP algorithm

In order to evaluate the performance of the WP algorithm, we use the first-order perturbation theory to prove the following theorem.

Theorem 1: Assume that $\mathbf{h}_1^{(1)}$ is estimated using (3.13). Then, the first-order perturbation theory-based approximation of the MSE of the estimation error $\delta\mathbf{h}_1^{(1)} = \hat{\mathbf{h}}_1^{(1)} - \mathbf{h}_1^{(1)}$ is given by

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1^{(1)}\|^2\right\} \approx \frac{1}{N}\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\mathbf{w}_1^{(1)^H}\mathbf{R}^{(12)^{\dagger^H}}\mathbf{R}^{(2)}\mathbf{R}^{(12)^\dagger}\mathbf{w}_1^{(1)} \tag{3.22}$$

where

$$\boldsymbol{\Psi} = \mathbf{U}_n^{(1)}\mathbf{T}_1^{(1)^{\dagger^H}}\mathbf{T}_1^{(1)^\dagger}\mathbf{U}_n^{(1)^H}. \tag{3.23}$$

Moreover, if the following conditions hold

$$\mathbf{w}_k^{(i)^H}\mathbf{w}_l^{(i)} = \|\mathbf{w}_k^{(i)}\|^2\delta_{kl}, \qquad i = 1,2 \tag{3.24}$$

$$\lambda_{\max}\big(\boldsymbol{\Sigma}_v^{(2)}\big) \ll \big(A_1^{(2)}\|\mathbf{w}_1^{(2)}\|\big)^2 \tag{3.25}$$

then (3.22) reduces to

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1^{(1)}\|^2\right\} \approx \frac{\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)}{N{A_1^{(1)}}^2}. \tag{3.26}$$

Proof: Since $\mathbf{U}_n^{(1)}$ spans the null-space of $\mathbf{R}^{(12)}$, we have

$$\mathbf{U}_n^{(1)^H}\mathbf{W}^{(1)} = \mathbf{0}. \tag{3.27}$$

Equations (3.27) and (3.23) yield

$$\mathbf{\Psi}\mathbf{W}^{(1)} = \mathbf{W}^{(1)^H}\mathbf{\Psi} = \mathbf{0}. \tag{3.28}$$

To prove (3.22), we introduce

$$\delta\mathbf{R}^{(12)} = \hat{\mathbf{R}}^{(12)} - \mathbf{R}^{(12)} \tag{3.29}$$

$$\delta\mathbf{U}_n^{(1)} = \hat{\mathbf{U}}_n^{(1)} - \mathbf{U}_n^{(1)}. \tag{3.30}$$

Using the perturbation theory, the first-order approximation of $\delta\mathbf{U}_n^{(1)}$ can be written as [40], [47]

$$\delta\mathbf{U}_n^{(1)} \doteq -\mathbf{R}^{(12)^{\dagger^H}}\delta\mathbf{R}^{(12)^H}\mathbf{U}_n^{(1)} \tag{3.31}$$

where

$$\mathbf{R}^{(12)^\dagger} = \mathbf{U}_s^{(2)}\mathbf{\Omega}_s^{(12)^{-1}}\mathbf{U}_s^{(1)^H}. \tag{3.32}$$

Since

$$\hat{\mathbf{U}}_n^{(1)^H}\boldsymbol{\mathcal{C}}_1\hat{\mathbf{h}}_1^{(1)} \approx \mathbf{0} \tag{3.33}$$

it follows that $\delta\mathbf{h}_1^{(1)}$ can be estimated as

$$\delta\mathbf{h}_1^{(1)} \approx -\mathbf{T}_1^{(1)^\dagger}\delta\mathbf{U}_n^{(1)^H}\mathbf{w}_1^{(1)}. \tag{3.34}$$

Inserting (3.31) into (3.34) and applying the expectation operation to the squared norm of the resulting expression, we have

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1^{(1)}\|^2\right\} \approx \mathbf{w}_1^{(1)^H}\mathbf{R}^{(12)^{\dagger^H}}\mathrm{E}\left\{\delta\mathbf{R}^{(12)^H}\mathbf{\Psi}\delta\mathbf{R}^{(12)}\right\}\mathbf{R}^{(12)^\dagger}\mathbf{w}_1^{(1)}. \tag{3.35}$$

Let us introduce

$$\mathbf{\Xi} = \mathrm{E}\left\{\delta\mathbf{R}^{(12)^H}\mathbf{\Psi}\delta\mathbf{R}^{(12)}\right\}. \tag{3.36}$$

From (3.28) and (3.29), it follows that

$$\mathbf{\Xi} = \mathrm{E}\left\{\hat{\mathbf{R}}^{(12)^H}\mathbf{\Psi}\hat{\mathbf{R}}^{(12)}\right\}. \tag{3.37}$$

Using (3.1) and (3.12) in (3.37) yields

$$\boldsymbol{\Xi} = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\mathrm{E}\left\{\left(\mathbf{W}^{(2)}\mathbf{b}(j)+\mathbf{v}^{(2)}(j)\right)\left(\mathbf{b}(j)^H\mathbf{W}^{(1)\,H}+\mathbf{v}^{(1)\,H}(j)\right)\boldsymbol{\Psi}\right.$$
$$\left.\left(\mathbf{W}^{(1)}\mathbf{b}(k)+\mathbf{v}^{(1)}(k)\right)\left(\mathbf{b}(k)^H\mathbf{W}^{(2)\,H}+\mathbf{v}^{(2)\,H}(k)\right)\right\}. \quad (3.38)$$

Using (3.28) to simplify the resulting expression, we obtain

$$\boldsymbol{\Xi} = \frac{1}{N}\left(\boldsymbol{\Phi}_1+\boldsymbol{\Phi}_2\right) \quad (3.39)$$

where

$$\boldsymbol{\Phi}_1 = \mathrm{E}\left\{\mathbf{W}^{(2)}\mathbf{b}(j)\mathbf{v}^{(1)\,H}(j)\boldsymbol{\Psi}\mathbf{v}^{(1)}(j)\mathbf{b}^H(j)\mathbf{W}^{(2)\,H}\right\} \quad (3.40)$$
$$\boldsymbol{\Phi}_2 = \mathrm{E}\left\{\mathbf{v}^{(2)}(j)\mathbf{v}^{(1)\,H}(j)\boldsymbol{\Psi}\mathbf{v}^{(1)}(j)\mathbf{v}^{(2)\,H}(j)\right\} \quad (3.41)$$

We have

$$\begin{aligned}\boldsymbol{\Phi}_1 &= \mathrm{E}\left\{\mathbf{v}^{(1)\,H}(j)\boldsymbol{\Psi}\mathbf{v}^{(1)}(j)\right\}\mathbf{W}^{(2)}\mathbf{W}^{(2)\,H}\\ &= \mathrm{E}\left\{\mathrm{tr}\left(\mathbf{v}^{(1)\,H}(j)\boldsymbol{\Psi}\mathbf{v}^{(1)}(j)\right)\right\}\mathbf{W}^{(2)}\mathbf{W}^{(2)\,H}\\ &= \mathrm{tr}\left(\mathrm{E}\left\{\mathbf{v}^{(1)}(j)\mathbf{v}^{(1)\,H}(j)\boldsymbol{\Psi}\right\}\right)\mathbf{W}^{(2)}\mathbf{W}^{(2)\,H}\\ &= \mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\mathbf{W}^{(2)}\mathbf{W}^{(2)\,H}. \quad (3.42)\\ \boldsymbol{\Phi}_2 &= \mathrm{E}\left\{\mathbf{v}^{(1)\,H}(j)\boldsymbol{\Psi}\mathbf{v}^{(1)}(j)\right\}\mathrm{E}\left\{\mathbf{v}^{(2)}(j)\mathbf{v}^{(2)\,H}(j)\right\}\\ &= \mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\boldsymbol{\Sigma}_v^{(2)}. \quad (3.43)\end{aligned}$$

Substituting (3.42) and (3.43) into (3.39) and using (3.4), we obtain

$$\boldsymbol{\Xi} = \frac{1}{N}\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\mathbf{R}^{(2)}. \quad (3.44)$$

Using (3.44) in (3.35) directly yields (3.22).

To prove (3.26), first let us use (3.6) and (3.24) to obtain

$$\mathbf{U}_s^{(1)} = \left[\frac{\mathbf{w}_1^{(1)}}{\|\mathbf{w}_1^{(1)}\|},\ldots,\frac{\mathbf{w}_K^{(1)}}{\|\mathbf{w}_K^{(1)}\|}\right] \quad (3.45)$$
$$\boldsymbol{\Omega}_s^{(12)} = \mathrm{diag}\left\{A_1^{(1)}A_1^{(2)}\|\mathbf{w}_1^{(1)}\|\|\mathbf{w}_1^{(2)}\|,\ldots,A_K^{(1)}A_K^{(2)}\|\mathbf{w}_K^{(1)}\|\|\mathbf{w}_K^{(2)}\|\right\} \quad (3.46)$$
$$\mathbf{U}_s^{(2)} = \left[\frac{\mathbf{w}_1^{(2)}}{\|\mathbf{w}_1^{(2)}\|},\ldots,\frac{\mathbf{w}_K^{(2)}}{\|\mathbf{w}_K^{(2)}\|}\right]. \quad (3.47)$$

Let

$$\breve{\mathbf{w}}_1^{(1)} = \mathbf{R}^{(12)^\dagger}\mathbf{w}_1^{(1)} = \mathbf{U}_s^{(2)}\mathbf{\Omega}_s^{(12)^{-1}}\mathbf{U}_s^{(1)^H}\mathbf{w}_1^{(1)}. \tag{3.48}$$

Substituting (3.45)-(3.47) into (3.48) and using (3.24), we have

$$\breve{\mathbf{w}}_1^{(1)} = \frac{1}{A_1^{(1)}A_1^{(2)}\|\mathbf{w}_1^{(2)}\|^2}\mathbf{w}_1^{(2)}. \tag{3.49}$$

Using (3.49) along with (3.22) yields

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1^{(1)}\|^2\right\} \approx \frac{\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)}{NA_1^{(1)^2}A_1^{(2)^2}\|\mathbf{w}_1^{(2)}\|^4}\mathbf{w}_1^{(2)^H}\mathbf{R}^{(2)}\mathbf{w}_1^{(2)}. \tag{3.50}$$

Substituting $\mathbf{R}^{(2)}$ from (3.4) into (3.50) and using (3.24) to simplify the result, we obtain

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1^{(1)}\|^2\right\} \approx \frac{\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)}{NA_1^{(1)^2}}\left(1 + \frac{\mathbf{w}_1^{(2)^H}\mathbf{\Sigma}_v^{(2)}\mathbf{w}_1^{(2)}}{A_1^{(2)^2}\|\mathbf{w}_1^{(2)}\|^4}\right). \tag{3.51}$$

As for any $\mathbf{w}_1^{(2)}$ and $\mathbf{\Sigma}_v^{(2)}$,

$$\mathbf{w}_1^{(2)^H}\mathbf{\Sigma}_v^{(2)}\mathbf{w}_1^{(2)} \le \|\mathbf{w}_1^{(2)}\|^2\lambda_{\max}(\mathbf{\Sigma}_v^{(2)}) \tag{3.52}$$

then, when (3.25) holds, (3.26) directly follows from (3.51). This completes the proof. □

Note that the average received power of the first user at the second antenna is equal to the right-hand side of (3.25), while the average noise power at the same antenna is lower bounded by the left-hand side because

$$\mathrm{E}\left\{\|\mathbf{v}^{(2)}(n)\|^2\right\} = \mathrm{tr}\big(\mathbf{\Sigma}_v^{(2)}\big) \ge \lambda_{\max}\big(\mathbf{\Sigma}_v^{(2)}\big). \tag{3.53}$$

Hence, if SNR at the second antenna is reasonably high, it is guaranteed that (3.25) holds. Using this observation along with the fact that (3.24) approximately holds in most practical scenarios, we can view (3.26) as a simple approximation of (3.22) in the high SNR regime. It explicitly clarifies the MSE of the estimated channel vector in terms of the *environmental parameters* such as the received power of the user-of-interest at the first antenna, the number of data samples, as well as the noise covariance matrix $\mathbf{\Sigma}_v^{(1)}$.

Note that both the MSE expressions (3.22) and (3.26) depend on $\mathbf{\Sigma}_v^{(1)}$ only through $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$. To study the parameters which have impact on the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$, we first note that if the channel vector is uniquely identifiable, then

$$\mathtt{rank}(\mathbf{T}_1^{(1)}) = \mathtt{rank}(\mathbf{\Psi}) = L - 1. \tag{3.54}$$

Moreover, we have[1]

$$\tau = \mathtt{dim}\{\mathtt{null}(\mathbf{\Psi})\} = L_c - L + 1 - \mathtt{rank}(\mathbf{\Psi}) = L_c - 2(L-1). \tag{3.55}$$

The effects of different parameters on the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$ are separately clarified in the following discussion.

Effect of L_c and L: As $\mathbf{\Sigma}_v^{(1)}$ and $\mathbf{\Psi}$ are positive (semi-)definite matrices, it follows that $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$ is real and nonnegative. Note that the projection of $\mathbf{\Sigma}_v^{(1)}$ onto $\mathtt{null}(\mathbf{\Psi})$ does not have any effect on the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$ which depends only on the projection of $\mathbf{\Sigma}_v^{(1)}$ onto $\mathtt{range}(\mathbf{\Psi})$. Therefore, the larger the projection of $\mathbf{\Sigma}_v^{(1)}$ onto $\mathtt{null}(\mathbf{\Psi})$, the smaller the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$. Using the latter fact, the effect of the spreading factor and the channel length on $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$, and, consequently, on the performance of the WP algorithm can be explained as follows. From (3.55) it can be observed that if either the spreading factor L_c increases or the channel length L decreases, then $\mathtt{dim}\{\mathtt{null}(\mathbf{\Psi})\}$ increases. In the latter case, the projection of the columns of $\mathbf{\Sigma}_v^{(1)}$ onto $\mathtt{null}(\mathbf{\Psi})$ becomes larger, and, therefore, their contribution to the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$ becomes smaller.

Effect of the eigenvectors of $\mathbf{\Sigma}_v^{(1)}$: The directions of the eigenvectors of $\mathbf{\Sigma}_v^{(1)}$ with respect to the eigenvectors of $\mathbf{\Psi}$ has considerable impact on the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$. To show this, let us eigendecompose $\mathbf{\Psi}$ as

$$\mathbf{\Psi} = \mathbf{\Pi}\mathbf{\Theta}\mathbf{\Pi}^H \tag{3.56}$$

where

$$\mathbf{\Pi} = [\boldsymbol{\pi}_1 \ \boldsymbol{\pi}_2 \ \cdots \ \boldsymbol{\pi}_{L-1}] \tag{3.57}$$

is an $L_c - L + 1 \times L - 1$ matrix whose columns are the orthonormal eigenvectors associated with the decreasingly-ordered positive eigenvalues of $\mathbf{\Psi}$ that are the diagonal elements of

[1] It should be noticed from (3.23) that $\mathtt{range}(\mathbf{W}^{(1)})$ is a $K-$dimensional subspace in $\mathtt{null}(\mathbf{\Psi})$.

$$\mathbf{\Theta} = \mathrm{diag}\left\{\theta_1, \theta_2, \ldots, \theta_{L-1}\right\}. \tag{3.58}$$

In contrary to $\mathtt{rank}(\mathbf{\Psi})$,

$$m = \mathtt{rank}(\mathbf{\Sigma}_v^{(1)}) \tag{3.59}$$

may not be known. In fact, $\mathtt{rank}(\mathbf{\Sigma}_v^{(1)})$ may vary from $m = 1$ for the case of coherent interference to $m = L_c - L + 1$ for the case of full-rank noise. Let us consider an arbitrary value of m and eigendecompose $\mathbf{\Sigma}_v^{(1)}$ as

$$\mathbf{\Sigma}_v^{(1)} = \mathbf{U}_v \mathbf{\Gamma}_v \mathbf{U}_v^H \tag{3.60}$$

where $\mathbf{U}_v$ is an $L_c - L + 1 \times m$ matrix whose orthonormal columns are the eigenvectors associated with the decreasingly-ordered positive eigenvalues of $\mathbf{\Sigma}_v^{(1)}$ which are the diagonal elements of

$$\mathbf{\Gamma}_v = \mathrm{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_m\}. \tag{3.61}$$

The value of $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$, and, hence, the MSE expressions (3.22) and (3.26) critically depend on the direction of the columns of $\mathbf{U}_v$ relative to the columns of $\mathbf{\Pi}$. To explain this fact, let us fix the matrix $\mathbf{\Gamma}_v$ and find the matrices $\mathbf{U}_{v\max}$ and $\mathbf{U}_{v\min}$ which maximize and minimize $\mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big)$, respectively. It can be shown [15], [34] that

$$\max_{\mathbf{U}_v} \left\{ \mathrm{tr}\big(\mathbf{\Sigma}_v^{(1)}\mathbf{\Psi}\big) \right\} = \sum_{i=1}^{\tau_1} \gamma_i \theta_i, \qquad \tau_1 = \min\{L-1, m\} \tag{3.62}$$

and $\mathbf{U}_{v\max}$ is given by

$$\mathbf{U}_{v\max} = \begin{cases} [\boldsymbol{\pi}_1 \ \boldsymbol{\pi}_2 \ \cdots \ \boldsymbol{\pi}_m], & \text{if} \quad m \le L-1 \\ \\ [\mathbf{\Pi} \ \mathbf{\Pi}_{m-L+1}^{\perp}], & \text{if} \quad m > L-1 \end{cases} \tag{3.63}$$

where $\mathbf{\Pi}_l^{\perp}$ is an $L_c - L + 1 \times l$ matrix whose $l \le \tau$ columns are *arbitrarily* chosen from a set of τ orthonormal vectors in $\mathtt{null}(\mathbf{\Psi})$. According to (3.63), for a fixed $\mathbf{\Gamma}_v$, the MSE expressions (3.22) and (3.26) are maximal if the first τ_1 columns of $\mathbf{U}_v$ and $\mathbf{\Pi}$ coincide. In

turn, we have [15], [34]

$$\min_{\mathbf{U}_v}\left\{\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\right\}=\begin{cases}0, & \text{if} \quad m\le\tau\\ \sum_{i=1}^{m-\tau}\gamma_{\tau+i}\theta_{L-i}, & \text{if} \quad m>\tau\end{cases} \tag{3.64}$$

and $\mathbf{U}_{v\min}$ is given by

$$\mathbf{U}_{v\min}=\begin{cases}\boldsymbol{\Pi}_m^{\perp}, & \text{if} \quad m\le\tau\\ [\boldsymbol{\Pi}_\tau^{\perp}\ \boldsymbol{\pi}_{L-1}\ \cdots\ \boldsymbol{\pi}_{L-(m-\tau)}], & \text{if} \quad m>\tau\end{cases}. \tag{3.65}$$

According to (3.65), the necessary condition to minimize the MSE expressions (3.22) and (3.26) is that the first $\tau_2=\min\{m,\tau\}$ columns of $\mathbf{U}_v$ are in $\mathtt{null}(\boldsymbol{\Psi})$. Note that the matrix

$$\boldsymbol{\Sigma}_{v\min}=\mathbf{U}_{v\min}\boldsymbol{\Gamma}_v\mathbf{U}_{v\min}^H \tag{3.66}$$

has the maximum projection onto $\mathtt{null}(\boldsymbol{\Psi})$, that is, the space spanned by the eigenvectors associated with the τ_2 largest eigenvalues of $\boldsymbol{\Sigma}_{v\min}$ is in $\mathtt{null}(\boldsymbol{\Psi})$.

Assuming that the average noise power at the first antenna is given by e_o, i.e.,

$$\mathrm{E}\left\{\|\mathbf{v}^{(1)}(n)\|^2\right\}=\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\big)=\sum_{i=1}^{m}\gamma_i=e_o \tag{3.67}$$

we can also obtain the extremal values of the MSE expressions (3.22) and (3.26) as follows. Since for any pair of positive (semi-)definite matrices $\boldsymbol{\Sigma}_v^{(1)}$ and $\boldsymbol{\Psi}$ we have [34]

$$\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\le\lambda_{\max}(\boldsymbol{\Psi})\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\big) \tag{3.68}$$

it directly follows that

$$\mathrm{tr}\big(\boldsymbol{\Sigma}_v^{(1)}\boldsymbol{\Psi}\big)\le\theta_1 e_o \tag{3.69}$$

where, assuming that the largest eigenvalue of $\boldsymbol{\Psi}$ is unique, (3.69) holds with equality if and only if

$$\boldsymbol{\Sigma}_v^{(1)}=e_o\boldsymbol{\pi}_1\boldsymbol{\pi}_1^H. \tag{3.70}$$

Moreover, it is obvious that among all noise covariance matrices with $\sum_{i=1}^{m}\gamma_i=e_o$, those in the form of

$$\boldsymbol{\Sigma}_v^{(1)}=\boldsymbol{\Pi}_m^{\perp}\boldsymbol{\Gamma}_v{\boldsymbol{\Pi}_m^{\perp}}^H \tag{3.71}$$

result in the MSE expressions (3.22) and (3.26) equal to zero. It is interesting to observe from (3.70) and (3.71) that, given the average noise power at the first antenna, both the maximal and the minimal values of the MSE of the channel vector estimate are obtained when the noise covariance matrix is rank deficient. As a rank deficient covariance matrix can be attributed to a narrow-band interference, it follows that the performance of the WP algorithm can be more sensitive to a narrow-band interference than a full-rank colored noise.

Now, let us consider two important particular scenarios in which the WP algorithm may be used and discuss the pertaining results.

White noise scenario: If the noise at the first antenna is white, i.e.,

$$\mathbf{\Sigma}_v^{(1)} = {\sigma^{(1)}}^2 \mathbf{I} \tag{3.72}$$

then (3.26) reduces to

$$\mathrm{E}\left\{ \|\delta \mathbf{h}_1^{(1)}\|^2 \right\} \approx \frac{{\sigma^{(1)}}^2 \|{\mathbf{T}^{(1)}}^\dagger\|_F^2}{N {A_1^{(1)}}^2} \tag{3.73}$$

which is equal to the derived MSE of the LX algorithm in (2.28). Hence, even though the WP algorithm is proposed to estimate the channel vector in the presence of unknown correlated noise, it is also applicable to the white noise scenario. In the latter case, the performance of the WP algorithm is identical to that of the LX algorithm.

Multiple antenna systems: It follows from (3.26) that if the SNR at the second antenna is high enough so that (3.25) holds, then the MSE of the channel vector estimate between the user-of-interest and the first antenna is independent from $\mathbf{\Sigma}_v^{(2)}$ and the received power of this user at the second antenna. Let us consider a receiver with $M > 2$ antennas which are spatially separated so that the noises between the first antenna and all the other antennas are uncorrelated. Moreover, assume that the SNR is high enough:

$$\lambda_{\max}\big(\mathbf{\Sigma}_v^{(i)}\big) \ll \big(A_1^{(i)} \|\mathbf{w}_1^{(i)}\|\big)^2, \qquad i = 2, \ldots, M \tag{3.74}$$

and that we aim to estimate the channel vector between the first user and the first antenna using the WP algorithm. Since this algorithm is based on processing of the data cross-correlation matrix between the first antenna and another well-separated *auxiliary* antenna,

we have to choose the auxiliary antenna among the $M-1$ available antennas. However, it directly follows from (3.26) that if the aforementioned assumptions hold, the performance of the channel vector estimate is insensitive to the choice of such an antenna, i.e., the auxiliary antenna can be chosen *arbitrarily*.

3.3.2 Performance analysis of the BP algorithm

The following theorem analyzes the performance of the BP algorithm.

Theorem 2: Assume that the channel vector is estimated using the BP algorithm. Then, the first-order perturbation theory-based approximation of the MSE of the estimation error $\delta\mathbf{h}_1 = \hat{\mathbf{h}}_1 - \mathbf{h}_1$ is given by

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{1}{N}\mathbf{w}_1^H\tilde{\mathbf{R}}^{\dagger^H}\left\{\mathrm{tr}\big(\boldsymbol{\Sigma}_v\tilde{\boldsymbol{\Psi}}\big)\mathbf{R}^T + (\boldsymbol{\Sigma}_v\tilde{\boldsymbol{\Psi}}\boldsymbol{\Sigma}_v)^T\right\}\tilde{\mathbf{R}}^{\dagger}\mathbf{w}_1 \tag{3.75}$$

where

$$\boldsymbol{\Sigma}_v = \mathrm{E}\left\{\mathbf{v}(n)\mathbf{v}(n)^H\right\} \tag{3.76}$$

$$\mathbf{R} = \mathbf{W}\mathbf{W}^H + \boldsymbol{\Sigma}_v \tag{3.77}$$

$$\tilde{\boldsymbol{\Psi}} = \tilde{\mathbf{U}}_n\tilde{\mathbf{T}}_1^{\dagger^H}\tilde{\mathbf{T}}_1^{\dagger}\tilde{\mathbf{U}}_n^H. \tag{3.78}$$

Moreover, if (2.23) holds and

$$\lambda_{\max}(\boldsymbol{\Sigma}_v) \ll A_1^2\|\mathbf{w}_1\|^2 \tag{3.79}$$

then (3.75) reduces to

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\mathrm{tr}\big(\boldsymbol{\Sigma}_v\tilde{\boldsymbol{\Psi}}\big)}{NA_1^2}. \tag{3.80}$$

Proof: According to (3.15), we have

$$\tilde{\mathbf{U}}_n^H\mathbf{W} = \mathbf{0}. \tag{3.81}$$

From (3.78) along with (3.81) it follows that

$$\tilde{\boldsymbol{\Psi}}\mathbf{W} = \mathbf{W}^H\tilde{\boldsymbol{\Psi}} = \mathbf{0}. \tag{3.82}$$

Using the procedure similar to that of the proof of Theorem 1, it can be readily shown that

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \mathbf{w}_1^H \tilde{\mathbf{R}}^{\dagger^H} \mathrm{E}\left\{\delta\tilde{\mathbf{R}}^H \tilde{\mathbf{\Psi}} \delta\tilde{\mathbf{R}}\right\} \tilde{\mathbf{R}}^{\dagger} \mathbf{w}_1 \tag{3.83}$$

where

$$\delta\tilde{\mathbf{R}} = \hat{\tilde{\mathbf{R}}} - \tilde{\mathbf{R}}. \tag{3.84}$$

Let us denote

$$\tilde{\mathbf{\Xi}} = \mathrm{E}\left\{\delta\tilde{\mathbf{R}}^H \tilde{\mathbf{\Psi}} \delta\tilde{\mathbf{R}}\right\}. \tag{3.85}$$

Substituting (3.84) into (3.85), and then using (3.82) to simplify the result, we have

$$\tilde{\mathbf{\Xi}} = \mathrm{E}\left\{\hat{\tilde{\mathbf{R}}}^H \tilde{\mathbf{\Psi}} \hat{\tilde{\mathbf{R}}}\right\}. \tag{3.86}$$

Expanding the right-hand side of (3.86) according to (3.21), and then using (3.82) to simplify the resulting expression, we obtain

$$\tilde{\mathbf{\Xi}} = \frac{1}{N}(\tilde{\mathbf{\Phi}}_1 + \tilde{\mathbf{\Phi}}_2) \tag{3.87}$$

where

$$\tilde{\mathbf{\Phi}}_1 = \mathrm{E}\left\{\mathbf{v}^H(i)\tilde{\mathbf{\Psi}}\mathbf{v}(i)\mathbf{W}^*\mathbf{b}(i)\mathbf{b}^T(i)\mathbf{W}^T\right\} \tag{3.88}$$

$$\tilde{\mathbf{\Phi}}_2 = \mathrm{E}\left\{\mathbf{v}^*(i)\mathbf{v}^H(i)\tilde{\mathbf{\Psi}}\mathbf{v}(i)\mathbf{v}^T(i)\right\}. \tag{3.89}$$

Since the transmitted symbols are drawn from the BPSK constellation, we have

$$\begin{aligned}\tilde{\mathbf{\Phi}}_1 &= \mathrm{E}\left\{\mathbf{v}^H(i)\tilde{\mathbf{\Psi}}\mathbf{v}(i)\right\}\mathbf{W}^*\mathbf{W}^T \\ &= \mathrm{tr}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\big)\mathbf{W}^*\mathbf{W}^T\end{aligned} \tag{3.90}$$

where the second line of (3.90) can be derived using the same steps as in (3.42). To obtain $\tilde{\mathbf{\Phi}}_2$, it can be easily shown from (3.89) that

$$[\tilde{\mathbf{\Phi}}_2]_{kl} = \sum_{g=1}^{L_c-L+1}\sum_{m=1}^{L_c-L+1} [\tilde{\mathbf{\Psi}}]_{gm}\mathrm{E}\left\{[\mathbf{v}]_k^*[\mathbf{v}]_g^*[\mathbf{v}]_m[\mathbf{v}]_l\right\} \tag{3.91}$$

where the time index i has been dropped from $\mathbf{v}(i)$ for the sake of simplicity. Since $\mathbf{v}$ is a multivariate circular Gaussian random vector, we have [30]

$$\begin{aligned}\mathrm{E}\left\{[\mathbf{v}]_k^*[\mathbf{v}]_g^*[\mathbf{v}]_m[\mathbf{v}]_l\right\} &= \mathrm{E}\left\{[\mathbf{v}]_m[\mathbf{v}]_k^*\right\}\mathrm{E}\left\{[\mathbf{v}]_l[\mathbf{v}]_g^*\right\} + \mathrm{E}\left\{[\mathbf{v}]_m[\mathbf{v}]_g^*\right\}\mathrm{E}\left\{[\mathbf{v}]_l[\mathbf{v}]_k^*\right\} \\ &= [\mathbf{\Sigma}_v]_{mk}[\mathbf{\Sigma}_v]_{lg} + [\mathbf{\Sigma}_v]_{mg}[\mathbf{\Sigma}_v]_{lk}. \end{aligned} \tag{3.92}$$

Substituting (3.92) into (3.91), we obtain

$$\begin{aligned}[\tilde{\mathbf{\Phi}}_2]_{kl} &= \sum_{g=1}^{L_c-L+1} [\mathbf{\Sigma}_v]_{lg}\big[\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big]_{gk} + \big[\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big]_{gg}[\mathbf{\Sigma}_v]_{lk} \\ &= \big[\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big]_{lk} + \mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)[\mathbf{\Sigma}_v]_{lk}. \end{aligned} \tag{3.93}$$

From (3.93) it directly follows that

$$\tilde{\mathbf{\Phi}}_2 = \big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)^T + \mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)\mathbf{\Sigma}_v^T. \tag{3.94}$$

Substituting (3.90) and (3.94) into (3.87) and using the resulting expression in (3.83), we obtain (3.75).

To prove (3.80), we note that (2.23) along with (3.15) yield

$$\tilde{\mathbf{U}}_s = \left[\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \ldots, \frac{\mathbf{w}_K}{\|\mathbf{w}_K\|}\right] \tag{3.95}$$

$$\tilde{\mathbf{\Omega}}_s = \mathrm{diag}\left\{A_1^2\|\mathbf{w}_1\|^2, \ldots, A_K^2\|\mathbf{w}_K\|^2\right\} \tag{3.96}$$

$$\tilde{\mathbf{V}}_s = \left[\frac{\mathbf{w}_1^*}{\|\mathbf{w}_1\|}, \ldots, \frac{\mathbf{w}_K^*}{\|\mathbf{w}_K\|}\right]. \tag{3.97}$$

Let us denote

$$\breve{\mathbf{w}}_1 = \tilde{\mathbf{R}}^{\dagger}\mathbf{w}_1 = \tilde{\mathbf{V}}_s\tilde{\mathbf{\Omega}}_s^{-1}\tilde{\mathbf{U}}_s^H\mathbf{w}_1. \tag{3.98}$$

Substituting (3.95)-(3.97) into the right-hand side of (3.98) and using (2.23), we have

$$\breve{\mathbf{w}}_1 = \frac{\mathbf{w}_1^*}{A_1^2\|\mathbf{w}_1\|^2}. \tag{3.99}$$

Using (3.99) in (3.75) results in the following expression for $\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\}$:

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}{N\|\mathbf{w}_1\|^2A_1^4}(\alpha_1+\alpha_2+\alpha_3) \tag{3.100}$$

where

$$\alpha_1 = \frac{\mathbf{w}_1^T}{\|\mathbf{w}_1\|}\mathbf{W}^*\mathbf{W}^T\frac{\mathbf{w}_1^*}{\|\mathbf{w}_1\|} \tag{3.101}$$

$$\alpha_2 = \frac{\mathbf{w}_1^T}{\|\mathbf{w}_1\|}\mathbf{\Sigma}_v^T\frac{\mathbf{w}_1^*}{\|\mathbf{w}_1\|} \tag{3.102}$$

$$\alpha_3 = \frac{1}{\mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}\frac{\mathbf{w}_1^T}{\|\mathbf{w}_1\|}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)^T\frac{\mathbf{w}_1^*}{\|\mathbf{w}_1\|}. \tag{3.103}$$

It directly follows from (2.23) that

$$\alpha_1 = A_1^2\|\mathbf{w}_1\|^2. \tag{3.104}$$

Noting that both $\mathbf{\Sigma}_v$ and $\tilde{\mathbf{\Psi}}$ are positive (semi-)definite matrices, it is easy to find an upper-bound for α_2 and α_3 as

$$\alpha_2 = \left(\frac{\mathbf{w}_1^H}{\|\mathbf{w}_1\|}\mathbf{\Sigma}_v\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right)^* \leq \lambda_{\max}^*(\mathbf{\Sigma}_v) = \lambda_{\max}(\mathbf{\Sigma}_v), \tag{3.105}$$

$$\begin{aligned}\alpha_3 &= \frac{1}{\mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}\left(\frac{\mathbf{w}_1^H}{\|\mathbf{w}_1\|}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right)^* \\ &\leq \frac{1}{\mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}\lambda_{\max}^*\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big) \\ &= \frac{1}{\mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}\lambda_{\max}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big) \\ &\leq \frac{\lambda_{\max}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}{\mathrm{tr}\big(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\big)}\lambda_{\max}(\mathbf{\Sigma}_v) \\ &\leq \lambda_{\max}(\mathbf{\Sigma}_v).\end{aligned} \tag{3.106}$$

Hence, if (3.79) holds, both α_2 and α_3 are negligible comparing to α_1. Substituting (3.104) into (3.100) directly yields (3.80). This completes the proof. □

As can be observed from (3.80), in the high SNR regime the MSE of the channel vector estimate of the BP algorithm can be expressed in terms of the noise covariance matrix, the received signal power, and the number of data samples.

Note that if the channel vector is uniquely identifiable from the BP algorithm, we have

$$\mathtt{rank}(\tilde{\mathbf{\Psi}}) = L - 1. \tag{3.107}$$

Comparing (3.80) with (3.26), it can be readily shown that the effect of the spreading factor and the channel length on both the WP and BP algorithms are similar. Moreover, following a discussion similar to that of Section 3.3.1, we can obtain the extremal values of $\mathrm{tr}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\big)$, and, consequently, those of the MSE expression (3.80). Let us first eigendecompose $\tilde{\mathbf{\Psi}}$ as

$$\tilde{\mathbf{\Psi}} = \tilde{\mathbf{\Pi}}\tilde{\mathbf{\Theta}}\tilde{\mathbf{\Pi}}^H \tag{3.108}$$

where

$$\tilde{\mathbf{\Pi}} = [\tilde{\boldsymbol{\pi}}_1\ \tilde{\boldsymbol{\pi}}_2\ \cdots\ \tilde{\boldsymbol{\pi}}_{L-1}] \tag{3.109}$$

contains the orthonormal eigenvectors associated with the positive eigenvalues of $\tilde{\mathbf{\Psi}}$ and

$$\tilde{\mathbf{\Theta}} = \mathrm{diag}\left\{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{L-1}\right\} \tag{3.110}$$

is the diagonal matrix that contains the decreasingly-ordered positive eigenvalues. Let us denote

$$q = \mathtt{rank}(\mathbf{\Sigma}_v) \tag{3.111}$$

and eigendecompose $\mathbf{\Sigma}_v$ as

$$\mathbf{\Sigma}_v = \tilde{\mathbf{U}}_v\tilde{\mathbf{\Gamma}}_v\tilde{\mathbf{U}}_v^H \tag{3.112}$$

where $\tilde{\mathbf{U}}_v$ contains the orthonormal eigenvectors associated with the positive eigenvalues of $\mathbf{\Sigma}_v$ which are ordered decreasingly as the diagonal elements of

$$\tilde{\mathbf{\Gamma}}_v = \mathrm{diag}\{\tilde{\gamma}_1, \tilde{\gamma}_2, \ldots, \tilde{\gamma}_q\}. \tag{3.113}$$

Denoting $\tilde{\mathbf{\Pi}}_l^{\perp}$ as an $L_c - L + 1 \times l$ matrix whose columns are orthonormal vectors in $\mathtt{null}(\tilde{\mathbf{\Psi}})$, we have

- For any given $\tilde{\mathbf{\Gamma}}_v$,

$$\max_{\tilde{\mathbf{U}}_v}\left\{\mathrm{tr}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\big)\right\} = \sum_{i=1}^{\tilde{\tau}_1}\tilde{\gamma}_i\tilde{\theta}_i, \qquad \tilde{\tau}_1 = \min\{L-1, q\} \tag{3.114}$$

  where the matrix $\tilde{\mathbf{U}}_v$ which maximizes $\mathrm{tr}\big(\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}\big)$ is given by

$$\tilde{\mathbf{U}}_{v\max} = \begin{cases} [\tilde{\boldsymbol{\pi}}_1\ \tilde{\boldsymbol{\pi}}_2\ \cdots\ \tilde{\boldsymbol{\pi}}_q], & \text{if} \quad q \le L-1 \\ \\ [\tilde{\mathbf{\Pi}}\ \tilde{\mathbf{\Pi}}_{q-L+1}^{\perp}], & \text{if} \quad q > L-1 \end{cases}. \tag{3.115}$$

- For any given $\tilde{\mathbf{\Gamma}}_v$,

$$\min_{\tilde{\mathbf{U}}_v} \left\{ \mathrm{tr}\big(\mathbf{\Sigma}_v \tilde{\mathbf{\Psi}}\big) \right\} = \begin{cases} 0, & \text{if} \quad q \leq \tau \\ \sum_{i=1}^{q-\tau} \tilde{\gamma}_{\tau+i} \tilde{\theta}_{L-i}, & \text{if} \quad q > \tau \end{cases} \tag{3.116}$$

where the matrix $\tilde{\mathbf{U}}_v$ which minimizes $\mathrm{tr}\big(\mathbf{\Sigma}_v \tilde{\mathbf{\Psi}}\big)$ is

$$\tilde{\mathbf{U}}_{v_{\min}} = \begin{cases} \tilde{\mathbf{\Pi}}_q^{\perp}, & \text{if} \quad q \leq \tau \\ [\tilde{\mathbf{\Pi}}_{\tau}^{\perp} \ \tilde{\boldsymbol{\pi}}_{L-1} \ \cdots \ \tilde{\boldsymbol{\pi}}_{L-(q-\tau)}], & \text{if} \quad q > \tau \end{cases} . \tag{3.117}$$

Comparing (3.114)-(3.117) with (3.62)-(3.65), it can be observed that the conclusions which follow from equations (3.62)-(3.65) can be easily extended to the BP algorithm, and, hence, we do not repeat them for the sake of brevity.

Let us also consider the case that the average noise power is given by e_o, i.e.,

$$\mathrm{tr}(\mathbf{\Sigma}_v) = \sum_{i=1}^{q} \tilde{\gamma}_i = e_o. \tag{3.118}$$

In such a case, assuming that the largest eigenvalue of $\tilde{\mathbf{\Psi}}$ is unique, the noise covariance matrix which maximizes $\mathrm{tr}\big(\mathbf{\Sigma}_v \tilde{\mathbf{\Psi}}\big)$ is given by

$$\mathbf{\Sigma}_v = e_o \tilde{\boldsymbol{\pi}}_1 \tilde{\boldsymbol{\pi}}_1^H. \tag{3.119}$$

Moreover, over all noise covariance matrices $\mathbf{\Sigma}_v$ with $\sum_{i=1}^{q} \tilde{\gamma}_i = e_o$, the value of $\mathrm{tr}\big(\mathbf{\Sigma}_v \tilde{\mathbf{\Psi}}\big)$ and, consequently, that of the MSE expression (3.80) is zero if and only if

$$\mathbf{\Sigma}_v = \tilde{\mathbf{\Pi}}_q^{\perp} \tilde{\mathbf{\Gamma}}_v \tilde{\mathbf{\Pi}}_q^{\perp H}. \tag{3.120}$$

Similar to the WP algorithm, it follows from (3.119) and (3.120) that the performance of the BP algorithm can be more sensitive to the narrow-band interference than to the full-rank noise.

If noise is white, i.e.,

$$\mathbf{\Sigma}_v = \sigma^2 \mathbf{I} \tag{3.121}$$

the MSE expression (3.80) reduces to

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\sigma^2\|\tilde{\mathbf{T}}_1^{\dagger}\|_F^2}{NA_1^2}. \tag{3.122}$$

Hence, the performances of the BP and the LX algorithm are identical in the white noise scenario. Therefore, the BP algorithm can also be applied to estimate the channel vector in the white noise case without any estimation performance loss as compared to the conventional LX algorithm.

Another interesting relationship between the WP and BP algorithms follows from the comparison of (3.26) and (3.80). Let the users transmit BPSK-modulated symbols and the receiver be equipped with two well-separated antennas such that the noise is spatially uncorrelated between them. Also, let (3.24) and (3.25) hold and

$$\lambda_{\max}\big(\mathbf{\Sigma}_v^{(1)}\big) \ll \big(A_1^{(1)}\|\mathbf{w}_1^{(1)}\|\big)^2. \tag{3.123}$$

Then, the MSE expressions (3.26) and (3.80) can be readily verified to coincide in the following two cases: when $\mathbf{h}_1^{(1)}$ is estimated using the WP algorithm with both antennas, and when $\mathbf{h}_1^{(1)}$ is estimated using the BP algorithm with only the first antenna.

3.4 Simulations

In this section, we validate our analytical results via computer simulations. In all examples, we consider $K = 7$ users that transmit BPSK-modulated symbols. Each point of the simulation curves is the result of averaging over 200 Monte-Carlo realizations of the noise and transmitted data sequences. In Figs. 3.1-3.8, Gold codes of length $L_c = 31$ are employed as the user spreading sequences and the channel vectors of length $L = 4$ are independently drawn from a zero-mean white complex Gaussian process and then scaled to become unit-norm vectors. In Figs. 3.1-3.5 and 3.9, the received noise at each antenna is considered to be a circular Gaussian process with $[\mathbf{\Sigma}_v]_{ij} = 0.7^{|i-j|}$. In the figures where the MSE of the channel estimate is drawn versus SNR, it is assumed that $N = 80$ data samples are used to estimate the channel.

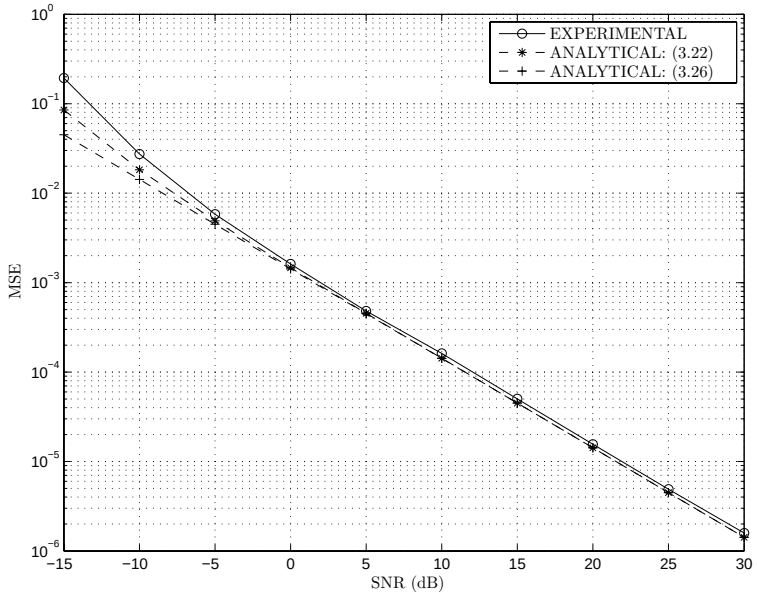

Figure 3.1: The MSE of the estimated channel versus SNR. The WP algorithm.

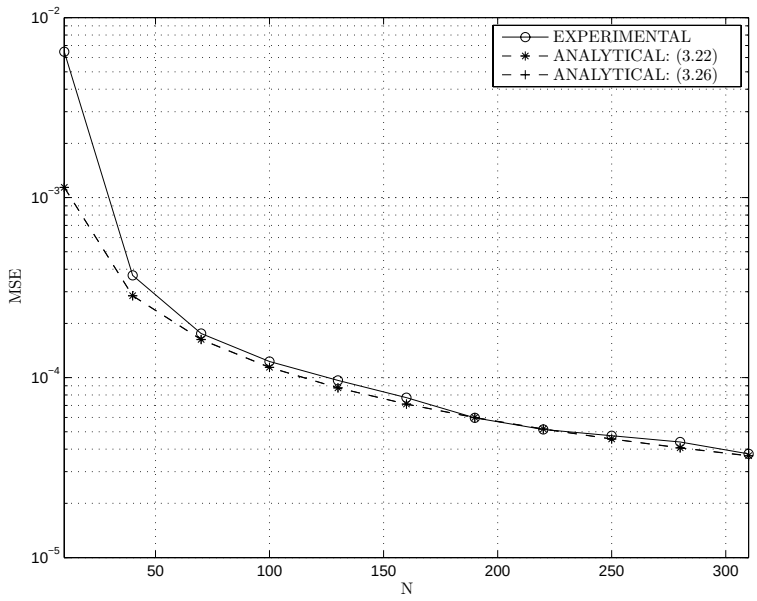

Figure 3.2: The MSE of the estimated channel versus number of data samples. The WP algorithm.

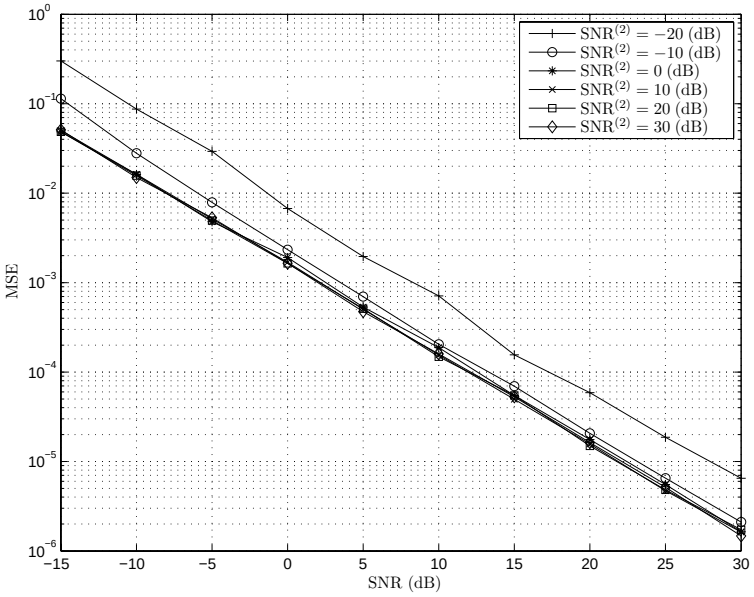

Figure 3.3: The MSE of the estimated channel versus SNR at the first antenna for different values of SNR at the second antenna. The WP algorithm.

Figs. 3.1-3.3 illustrate the accuracy of our analytical results derived for the WP algorithm. In Figure 3.1, it is assumed that SNRs of all users at both antennas are identical and $\mathbf{h}_1^{(1)}$ is estimated according to (3.13). The solid curve represents the MSE resulting from this estimate. This curve is compared with our analytical results given by (3.22) and (3.26). It can be observed that both theoretical curves follow the experimental MSE curve with a good precision. Note that when the SNR is very low, the channel vector estimation error is quite large and, hence, it could not be reliably predicted using the first-order perturbation theory. In such a case, the analytical MSE curves obtained from (3.22) and (3.26) show considerable discrepancy with the experimental MSE curve.

Fig. 3.2 depicts the experimental and the analytical MSE curves versus the number of data samples N. In this figure, it is assumed that the received signal power from each user at each of the two antennas is equal to 10 dB. Due to the fact that the SNR is reasonably high, the theoretical curve (3.22) and its high SNR approximation (3.26) are almost indistinguishable from each other and follow the experimental MSE curve with a good accuracy. It can be observed from Figure 3.2 that, when the number of data samples

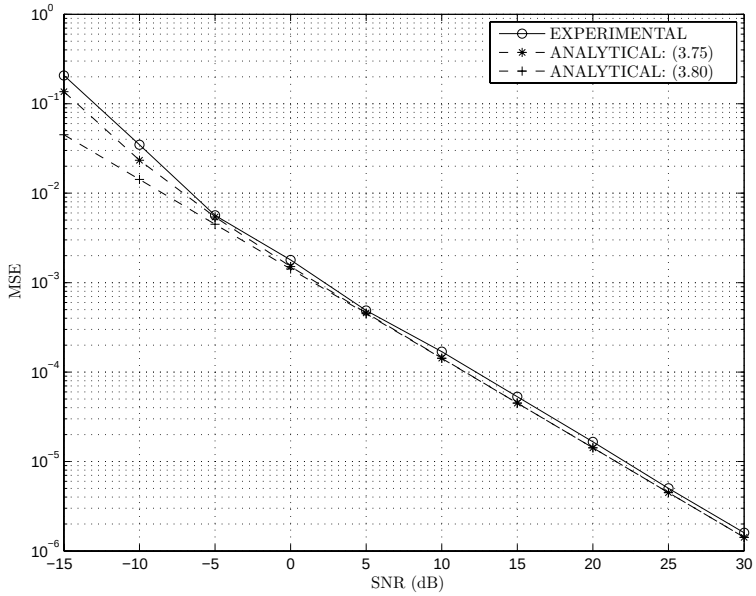

Figure 3.4: The MSE of the estimated channel versus SNR. The BP algorithm.

N is small, the small perturbation assumption is violated, and, hence, the accuracy of the analytical MSE curves decreases.

Fig. 3.3 shows the MSE of estimated channel $\hat{\mathbf{h}}_1^{(1)}$ versus SNR at the first antenna ($\text{SNR}^{(1)}$) for six different values of SNR at the second antenna ($\text{SNR}^{(2)}$). As expected from Section 3.3.1, the performance of the channel estimation is almost independent from the exact value of $\text{SNR}^{(2)}$, unless $\text{SNR}^{(2)}$ is very low.

Figs. 3.4-3.7 and 3.9 show the performance of the BP algorithm and compare it to our analytical results. In Figure 3.4, the experimental MSE curve is plotted versus SNR and is compared with the theoretical curves obtained from (3.75) and (3.80). As can be observed from the figure, the two theoretical MSE curves are very close to each other and also closely follow the experimental MSE curve for the SNRs higher than 0 dB.

Fig. 3.5 shows the experimental and the theoretical curves drawn versus the number of data samples N for SNR = 10 dB. As the figure shows, the theoretical curve (3.75) is precisely followed by its high SNR approximation (3.80) and both of them are very close to the experimental MSE curve.

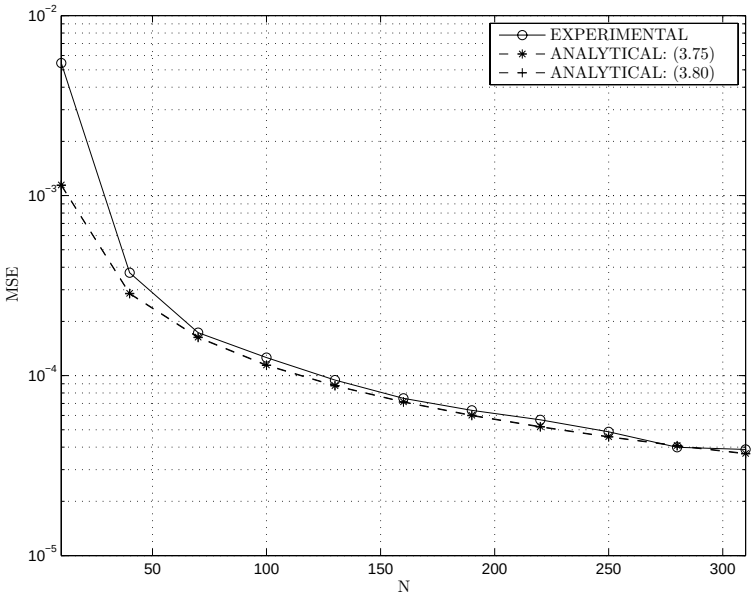

Figure 3.5: The MSE of the estimated channel versus number of data samples. The BP algorithm.

Fig. 3.6 shows the experimental MSE curves versus SNR for noise covariance matrices with the same $\tilde{\mathbf{\Gamma}}_v = \text{diag}\{20, 5, 3\}$ and different matrices of eigenvectors $\tilde{\mathbf{U}}_v$. Three random realizations of $\tilde{\mathbf{U}}_v$ as well as $\tilde{\mathbf{U}}_{v_{\max}}$ and $\tilde{\mathbf{U}}_{v_{\min}}$ are drawn and then, using (3.112), the corresponding noise covariance matrices are obtained. The BP algorithm is used to estimate the channel vector in the presence of a correlated noise with the so-obtained noise covariance matrices. Fig. 3.6 confirms our theoretical results in Section 3.3.2 which state that the worst and the best MSE performances are obtained when $\tilde{\mathbf{U}}_v = \tilde{\mathbf{U}}_{v_{\max}}$ and $\tilde{\mathbf{U}}_v = \tilde{\mathbf{U}}_{v_{\min}}$, respectively. Note that if $\tilde{\mathbf{U}}_v = \tilde{\mathbf{U}}_{v_{\min}}$, then, unlike the MSE expression (3.80), the experimental MSE performance is not equal to zero. This is due to the fact that the MSE expression (3.80) is obtained using the first-order perturbation theory and even in high SNR regime this expression has a slight difference with the experimental MSE.

Fig. 3.7 plots the experimental MSE curves versus SNR for several types of noise with the same average energy of $e_o = L_c - L + 1 = 28$ and different covariance matrices. For each value of q ($q = 1,\ 5,\ 15$), one noise covariance matrix is drawn randomly and another one is obtained according to (3.120). A rank-one noise covariance matrix is also derived

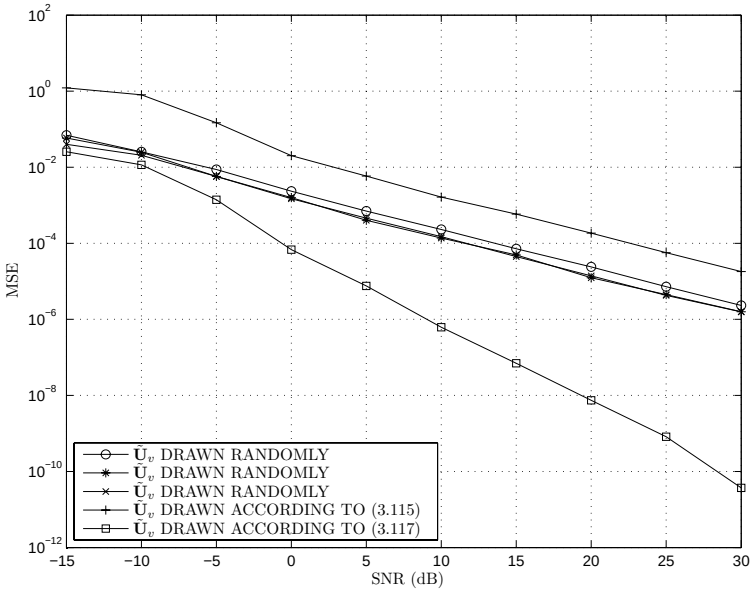

Figure 3.6: The MSE of the estimated channel versus SNR for $\tilde{\mathbf{\Gamma}}_v = \text{diag}\{20, 5, 3\}$ and different matrices $\tilde{\mathbf{U}}_v$. The BP algorithm.

according to (3.119). Then, the BP algorithm is used to estimate the channel vector in the presence of correlated noise with the so-obtained noise covariance matrices. Our analytical results in Section 3.3.2 are validated by means of observing that the worst and the best MSE performances are obtained when the noise covariance matrix follows (3.119) and (3.120), respectively.

In Fig. 3.8, the performance of the LX, WP, and BP algorithms are tested in the white noise environment. As predicted by our analysis in Section 2.5, all three methods have nearly identical performance.

Fig. 3.9 shows the experimental and the theoretical MSE curves of the BP algorithm versus the channel length L for two different values of the spreading factor $L_c = 40$ and $L_c = 80$. In this example, we use random spreading codes rather than the optimized Gold codes. The entries of these codes are randomly drawn from the set $\{-1, +1\}$. From Fig. 3.9 we see that, as predicted in Section 2.5, the estimation performance decreases with increasing L. When $L_c = 80$, the MSE of the channel vector estimate is significantly lower than that for $L_c = 40$. It can be observed that the curves corresponding to (3.75) and (3.80)

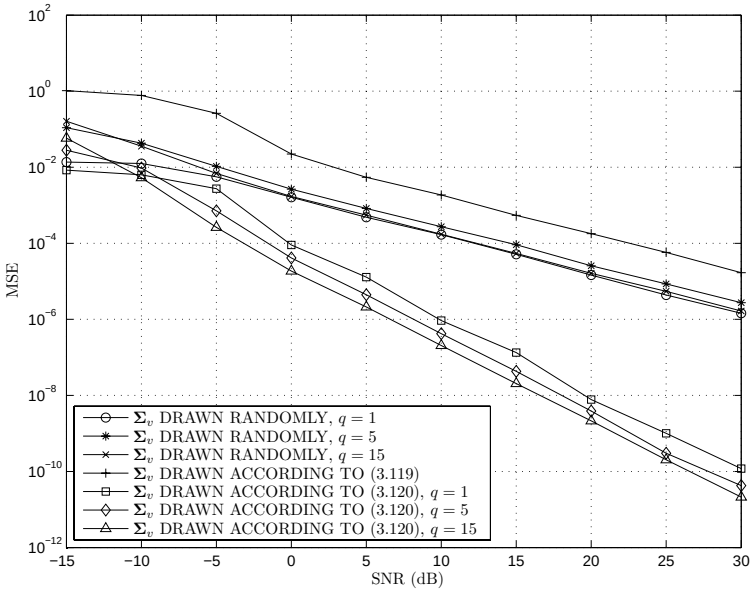

Figure 3.7: MSEs of the estimated channel versus SNR for $e_o = 28$ and different matrices $\mathbf{\Sigma}_v$. The BP algorithm.

are quite close to each other and, therefore, the use of the random spreading codes instead of the Gold codes retains the accuracy of (3.80).

3.5 Chapter summary

In this chapter, analytical expressions for the MSE of the channel vector estimate of the WP and BP algorithms have been derived. Assuming that different user signature vectors are orthogonal, the simplified versions of these expressions have been also obtained for the high SNR regime. The effect of the correlated noise on the performance of both algorithms has been studied. It has been shown that the direction of the eigenvectors of the noise covariance matrix has a significant effect on the performance of both algorithms. In particular, assuming that the eigenvalues of the noise covariance matrix are fixed, the eigenvectors of this matrix corresponding to the extremal values of the MSEs have been obtained. Over all noise covariance matrices with identical average noise power, the extremal values of the

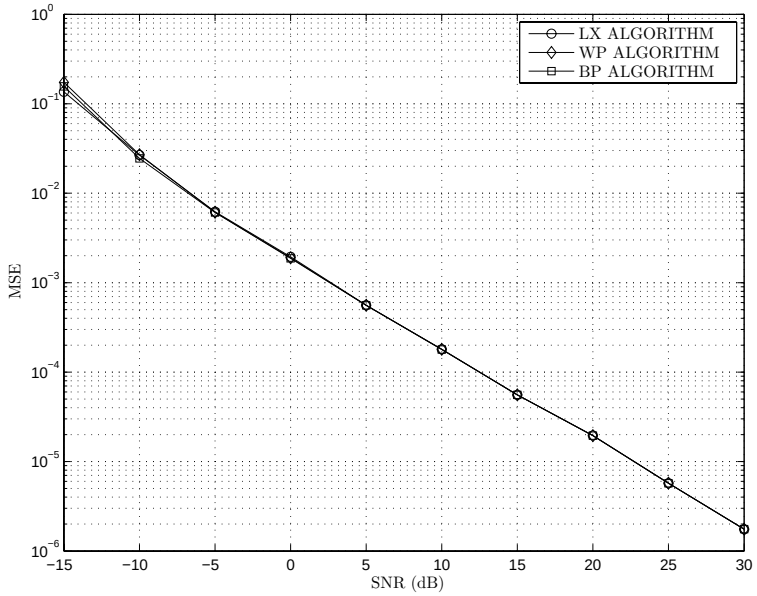

Figure 3.8: MSEs of the estimated channel versus SNR in the white noise environment. The LX, WP, and BP algorithms.

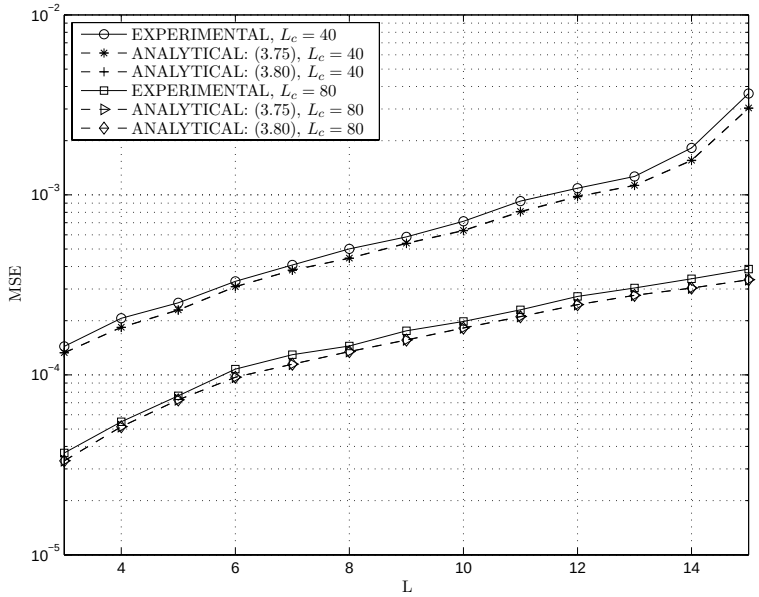

Figure 3.9: MSEs of the estimated channel versus L for $L_c = 40$ and $L_c = 80$. The BP algorithm.

MSEs have been derived and it has been shown that both the maximal and the minimal values of the MSEs are achieved when the noise covariance matrix is rank deficient. Moreover, it has been shown that at high SNRs and in the presence of white noise, the performance of the WP and BP techniques is identical to that of the conventional LX algorithm.

In the high SNR regime, it has been proved that the performance of the WP algorithm is independent from the noise covariance matrix and the user received power at the second auxiliary antenna. This property has been generalized to the multiple antenna systems and it has been shown that for such systems the choice of the auxiliary antenna is arbitrary at high SNRs.

Chapter 4

Signature estimation in the presence of wide-sense stationary noise

4.1 Introduction

As discussed in Chapter 3, the WP and BP algorithms are two key subspace-based signature estimation techniques in the presence of unknown correlated noise. In the WP algorithm, it is assumed that the receiver is equipped with two well-separated antennas such that the noise is spatially white between them, and the signal subspace is identified from the cross-correlation between the received data of these antennas [90]. Due to the fact that using two well-separated antennas in the current mobile receivers may be impractical, the BP algorithm [11] has been proposed to identify the signal subspace using just a single antenna at the receiver. This algorithm is based on the assumptions that the noise is a circular Gaussian random process and the user transmitted symbols are drawn from the BPSK constellation. Note that most of the leading standard proposals for the 3G wireless communication systems recommend symmetric constellations such as QPSK [9],

[81]. Therefore, practical applications of the latter technique may be limited.

In this chapter (see also [110], [116]), we propose an alternative approach to the problem of blind subspace-based signature waveform estimation in the presence of unknown correlated noise. Similar to [11], the proposed technique can be applied to a single-antenna receiver, but, in contrast to [11], it is applicable to the case of an arbitrary transmitted symbol constellation. In our technique, it is only assumed that the unknown correlated noise is wide-sense stationary. Note that this assumption has been frequently used in various interference rejection schemes for CDMA communication systems [11], [42], [51], [62], [88], and includes several particular interference models such as multi-tone and AR interference models [3], [20], [21], [39], [45], [67], [74], [105]. Both these models are widely used in the literature, in particular, to represent overlaid narrow-band interference whose current value can be effectively estimated from its past values [8], [105]. Exploiting the idea presented in [63] in the context of array processing, we obtain a subspace which is orthogonal to the subspace spanned by the user signals. We then use the so-obtained subspace along with the known spreading sequence of the user-of-interest to identify the channel vector, and, subsequently, the signature vector of this user. We also derive the necessary and sufficient conditions which warrant the identifiability of the proposed technique. Similar to other subspace-based signature estimation techniques [6], [47], [76], [89], [91], [101], the proposed algorithm poses some restrictions on the maximum admissible number of active users as well as on the length of the channel vector. We show that using the temporal oversampling technique, our algorithm can also be effectively applied to overloaded scenarios with lengthy channels.

Similar to other subspace-based estimation techniques, the channel vector can be precisely identified only if the exact data covariance matrix is available. We derive a closed-form analytical expression for the MSE of the estimated channel vector for the case that the data covariance matrix is estimated from the data samples. Then, a simplified version of this expression is also presented for the high SNR regime. From the latter expression, an impact of the key physical parameters (such as the number of data samples, the received power of the user-of-interest, and the received interference power) on the performance of the proposed

algorithm is studied.

The rest of this chapter is organized as follows. Our technique is presented in Section 4.2, where the necessary and sufficient identifiability conditions are also derived. The finite-sample performance of the proposed algorithm is analyzed in Section 4.3. In Section 4.4, the temporally oversampled version of the proposed technique is presented. Section 4.5 contains computer simulation results. Conclusions are drawn in Section 4.6.

4.2 The proposed technique

First note from (2.6) that the covariance matrix of the ISI-free part of the received data is given by

$$\mathbf{R} = \mathrm{E}\{\mathbf{x}(n)\mathbf{x}(n)^H\} = \mathbf{W}\mathbf{W}^H + \mathbf{\Sigma}_v \tag{4.1}$$

where $\mathbf{\Sigma}_v = \mathrm{E}\{\mathbf{v}(n)\mathbf{v}(n)^H\}$. Since $\mathbf{v}(n)$ is wide-sense stationary, the entries of $\mathbf{\Sigma}_v$ depend only on the difference between the observation times. Hence, $\mathbf{\Sigma}_v$ is a Hermitian Toeplitz matrix, and, therefore, it is centro-Hermitian [30], [63], that is,

$$\mathbf{J}\mathbf{\Sigma}_v^*\mathbf{J} = \mathbf{\Sigma}_v. \tag{4.2}$$

Using the idea presented in [63] for DOA estimation, let us exploit (4.2) to facilitate estimation of the user signature waveforms without knowing the correlation matrix $\mathbf{\Sigma}_v$. Let us form the covariance difference matrix [63] as follows:

$$\mathbf{R}_d = \mathbf{R} - \mathbf{J}\mathbf{R}^*\mathbf{J}. \tag{4.3}$$

From (4.1) and (4.2), it follows that

$$\mathbf{R}_d = \mathbf{W}\mathbf{W}^H - \mathbf{J}\mathbf{W}^*\mathbf{W}^T\mathbf{J}. \tag{4.4}$$

Note that $\mathbf{R}_d$ depends on the user transmitted signals while it is independent from the unknown interference covariance matrix $\mathbf{\Sigma}_v$. We can rewrite (4.4) as [63]

$$\mathbf{R}_d = \begin{bmatrix} \mathbf{W} & \mathbf{J}\mathbf{W}^* \end{bmatrix} \begin{bmatrix} \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_K \end{bmatrix} \begin{bmatrix} \mathbf{W} & \mathbf{J}\mathbf{W}^* \end{bmatrix}^H. \tag{4.5}$$

Since $\left[\begin{array}{cc} \mathbf{W} & \mathbf{JW}^* \end{array}\right]$ is an $(L_c - L + 1) \times 2K$ matrix, we have that if

$$L_c > 2K + L - 1 \tag{4.6}$$

then $\mathbf{R}_d$ is rank-deficient. Considering hereafter such a case, the matrix $\mathbf{R}_d$ can be eigen-decomposed as

$$\mathbf{R}_d = \left[\begin{array}{cc} \mathbf{U}_s & \mathbf{U}_n \end{array}\right] \left[\begin{array}{cc} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{array}\right] \tag{4.7}$$

where $\mathbf{\Lambda}_s$ is the $2K \times 2K$ diagonal matrix whose diagonal elements are the non-zero eigenvalues of $\mathbf{R}_d$ and $\mathbf{U}_s$ is the $(L_c - L + 1) \times 2K$ matrix whose columns are the eigenvectors associated with these eigenvalues. In turn, $\mathbf{U}_n$ is the $(L_c - L + 1) \times (L_c - L + 1 - 2K)$ matrix whose columns are the eigenvectors associated with the zero eigenvalues of $\mathbf{R}_d$.

The following lemma is essential for our later analysis.

Lemma 1: Assume that $\lambda > 0$ is an eigenvalue of $\mathbf{R}_d$ and $\mathbf{u}$ is its associated eigenvector. Then, $-\lambda$ is also an eigenvalue of this matrix and its associated eigenvector is $\mathbf{Ju}^*$. Moreover, there exists a unitary matrix $\mathbf{\Omega}$ such that

$$\mathbf{U}_n\mathbf{\Omega} = \mathbf{JU}_n^*. \tag{4.8}$$

Proof: The first part of this lemma has been proved in [63]. To prove (4.8), note from (4.3) that

$$\mathbf{JR}_d = -\mathbf{R}_d^*\mathbf{J}. \tag{4.9}$$

To prove (4.8), note that

$$\mathbf{R}_d\mathbf{U}_n = \mathbf{0}. \tag{4.10}$$

Left-multiplying of (4.10) with $\mathbf{J}$, using (4.9), and applying conjugate operation, we obtain

$$\mathbf{R}_d\mathbf{JU}_n^* = \mathbf{0}. \tag{4.11}$$

As columns of $\mathbf{U}_n$ span the null-space of $\mathbf{R}_d$, all of the columns of $\mathbf{JU}_n^*$ are in $\mathtt{range}(\mathbf{U}_n)$, or, equivalently, there exists a matrix $\mathbf{\Omega}$ such that

$$\mathbf{U}_n\mathbf{\Omega} = \mathbf{JU}_n^*. \tag{4.12}$$

Applying Hermitian operation to both sides of (4.12), we have

$$\mathbf{\Omega}^H \mathbf{U}_n^H = \mathbf{U}_n^T \mathbf{J}. \tag{4.13}$$

Left-multiplying each side of (4.13) with the corresponding side of (4.12) and using the fact that

$$\mathbf{U}_n^H \mathbf{U}_n = \mathbf{J}^2 = \mathbf{I} \tag{4.14}$$

we obtain

$$\mathbf{\Omega}^H \mathbf{\Omega} = \mathbf{I}. \tag{4.15}$$

Since $\mathbf{\Omega}$ is a square matrix, from (4.15) it follows that $\mathbf{\Omega}$ is unitary. This completes the proof. □

From (4.5) and (4.7) it follows that

$$\texttt{range}(\mathbf{R}_d) = \texttt{range}\left(\begin{bmatrix} \mathbf{W} & \mathbf{J}\mathbf{W}^* \end{bmatrix}\right) = \texttt{range}(\mathbf{U}_s). \tag{4.16}$$

Since all columns of $\mathbf{U}_n$ are orthogonal to all vectors in $\texttt{range}(\mathbf{U}_s)$, we have

$$\mathbf{U}_n^H \begin{bmatrix} \mathbf{W} & \mathbf{J}\mathbf{W}^* \end{bmatrix} = \mathbf{0}. \tag{4.17}$$

Let us recall that the first user is the user-of-interest. From (4.17), it follows that

$$\mathbf{U}_n^H \mathbf{w}_1 = \mathbf{0} \tag{4.18}$$

$$\mathbf{U}_n^H \mathbf{J}\mathbf{w}_1^* = \mathbf{0}. \tag{4.19}$$

It can be shown that (4.18) and (4.19) are equivalent, i.e., from either of them the other can be obtained. For instance, from (4.18) we have that $\mathbf{\Omega}^H \mathbf{U}_n^H \mathbf{w}_1 = \mathbf{0}$. Substituting (4.8) into the latter equation and applying conjugate operation, equation (4.19) follows. Equivalency of (4.18) and (4.19) shows that either (4.18) or (4.19) can be exploited to identify the channel vector $\mathbf{h}_1$. From (4.18) along with (2.5), it follows that

$$\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_1 \mathbf{h}_1 = \mathbf{0} \tag{4.20}$$

and, therefore, $\mathbf{h}_1$ is a nontrivial solution to

$$\mathbf{T}\mathbf{h} = \mathbf{0} \tag{4.21}$$

where

$$\mathbf{T} = \mathbf{U}_n^H \boldsymbol{\mathcal{C}}_1. \tag{4.22}$$

It is easy to verify that (4.21) is a linear system with $L_c - L + 1 - 2K$ equations and L unknowns. To have a unique nontrivial solution for (4.21), it is necessary that the number of equations is greater than or equal to the number of unknowns, that is,

$$L_c + 1 \geq 2K + 2L. \tag{4.23}$$

Obviously, (4.23) implies (4.6).

Equation (4.23) represents only the necessary condition of uniqueness of the solution to (4.21). The *necessary and sufficient* conditions for unique identifiability of $\mathbf{h}_1$ can be obtained as follows. As $\mathbf{h}_1$ is a nontrivial solution to (4.21), it follows that $\mathbf{w}_1 = \boldsymbol{\mathcal{C}}_1\mathbf{h}_1$ lies in both $\texttt{range}(\boldsymbol{\mathcal{C}}_1)$ and $\texttt{null}(\mathbf{U}_n) = \texttt{range}\left(\begin{bmatrix} \mathbf{W} & \mathbf{J}\mathbf{W}^* \end{bmatrix}\right)$. Therefore, $\mathbf{w}_1$ is in the intersection between the two latter subspaces and can be uniquely identified from (4.21) if and only if

$$\texttt{dim}\left\{\texttt{range}(\boldsymbol{\mathcal{C}}_1) \cap \texttt{range}\left(\begin{bmatrix} \mathbf{W} & \mathbf{J}\mathbf{W}^* \end{bmatrix}\right)\right\} = 1. \tag{4.24}$$

If we further assume that $\boldsymbol{\mathcal{C}}_1$ is a full column-rank matrix, then $\mathbf{w}_1$ corresponds to a unique channel vector $\mathbf{h}_1$. Hence, $\mathbf{h}_1$ can be uniquely identified from (4.21) if and only if (4.24) holds true and $\boldsymbol{\mathcal{C}}_1$ is full column-rank.

In practice, $\mathbf{R}$ is estimated from data samples as

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}(n)\mathbf{x}^H(n) \tag{4.25}$$

and the proposed method can be formulated as follows.

1. Compute the eigendecomposition of

$$\hat{\mathbf{R}}_d = \hat{\mathbf{R}} - \mathbf{J}\hat{\mathbf{R}}^*\mathbf{J} \tag{4.26}$$

as

$$\hat{\mathbf{R}}_d = \begin{bmatrix} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Lambda}}_s & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Lambda}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{bmatrix} \tag{4.27}$$

where the matrices $\hat{\mathbf{U}}_s$, $\hat{\mathbf{U}}_n$, and $\hat{\mathbf{\Lambda}}_s$ are the finite-sample estimates of the matrices $\mathbf{U}_s$, $\mathbf{U}_n$, and $\mathbf{\Lambda}_s$, respectively, and $\hat{\mathbf{\Lambda}}_n$ is the diagonal matrix whose diagonal elements are the $L_c - L + 1 - 2K$ eigenvalues of $\hat{\mathbf{R}}_d$ with the least absolute values.

2. Compute

$$\hat{\mathbf{T}} = \hat{\mathbf{U}}_n^H \boldsymbol{\mathcal{C}}_1 \tag{4.28}$$

and find the LS estimate of the channel $\mathbf{h}_1$ as

$$\hat{\mathbf{h}}_1 = \mho(\hat{\mathbf{T}}^H \hat{\mathbf{T}}). \tag{4.29}$$

4.3 Performance analysis

Using the first-order perturbation analysis, let us derive an approximate expression for the MSE of the channel vector estimate $\hat{\mathbf{h}}_1$. The following theorem holds.

Theorem 1: Assume that $\mathbf{h}_1$ is estimated using (4.29). Then, the first-order perturbation theory-based approximation of the MSE of the estimation error $\delta\mathbf{h}_1 = \hat{\mathbf{h}}_1 - \mathbf{h}_1$ is given by

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{1}{N} \mathbf{w}_1^H \mathbf{R}_d^{\dagger} \big(\mathrm{tr}\left(\mathbf{\Sigma}_v \mathbf{\Psi}\right)\left(\mathbf{W}\mathbf{W}^H + \mathbf{J}\mathbf{W}^*\mathbf{W}^T\mathbf{J}\right) + \mathbf{\Theta}\big) \mathbf{R}_d^{\dagger} \mathbf{w}_1 \tag{4.30}$$

where

$$\mathbf{\Psi} = \mathbf{U}_n \mathbf{T}^{\dagger H} \mathbf{T}^{\dagger} \mathbf{U}_n^H \tag{4.31}$$

$$\mathbf{\Theta} = \mathrm{E}\left\{(\mathbf{v}\mathbf{v}^H - \mathbf{J}\mathbf{v}^*\mathbf{v}^T\mathbf{J})\mathbf{\Psi}(\mathbf{v}\mathbf{v}^H - \mathbf{J}\mathbf{v}^*\mathbf{v}^T\mathbf{J})\right\}. \tag{4.32}$$

Note that for the sake of simplicity, the time index n of $\mathbf{i}(n)$ has been omitted in (4.32).

Proof: From (4.17) and (4.31), we have

$$\mathbf{\Psi}\mathbf{W} = \mathbf{\Psi}\mathbf{J}\mathbf{W}^* = \mathbf{0}. \tag{4.33}$$

It is also obvious that

$$\mathbf{\Psi}\mathbf{R}_d = \mathbf{0}. \tag{4.34}$$

Let us denote

$$\delta\mathbf{R}_d = \hat{\mathbf{R}}_d - \mathbf{R}_d \tag{4.35}$$

$$\delta\mathbf{U}_n = \hat{\mathbf{U}}_n - \mathbf{U}_n\,. \tag{4.36}$$

Using the perturbation theory, the first-order approximation of $\delta\mathbf{U}_n$ can be written as [40], [47], [99]

$$\delta\mathbf{U}_n \doteq -\mathbf{R}_d^{\dagger}\delta\mathbf{R}_d\mathbf{U}_n \tag{4.37}$$

where

$$\mathbf{R}_d^{\dagger} = \mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H. \tag{4.38}$$

As $\hat{\mathbf{U}}_n^H\boldsymbol{\mathcal{C}}_1\hat{\mathbf{h}}_1 \approx \mathbf{0}$, it follows that

$$\delta\mathbf{h}_1 \approx -\mathbf{T}^{\dagger}\delta\mathbf{U}_n^H\mathbf{w}_1\,. \tag{4.39}$$

Inserting (4.37) in (4.39) and applying the expectation operation to the squared norm of the resulting expression, we obtain

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \mathbf{w}_1^H\mathbf{R}_d^{\dagger}\ \mathrm{E}\left\{\delta\mathbf{R}_d\mathbf{\Psi}\delta\mathbf{R}_d\right\}\ \ \mathbf{R}_d^{\dagger}\mathbf{w}_1\,. \tag{4.40}$$

Hence, to find the MSE we need to obtain $\mathrm{E}\left\{\delta\mathbf{R}_d\mathbf{\Psi}\delta\mathbf{R}_d\right\}$. Note that

$$\begin{aligned}\mathrm{E}\left\{\delta\mathbf{R}_d\mathbf{\Psi}\delta\mathbf{R}_d\right\} &= \mathrm{E}\big\{\big(\hat{\mathbf{R}}_d - \mathbf{R}_d\big)\mathbf{\Psi}\big(\hat{\mathbf{R}}_d - \mathbf{R}_d\big)\big\}\\ &= \mathrm{E}\big\{\hat{\mathbf{R}}_d\mathbf{\Psi}\hat{\mathbf{R}}_d\big\}\end{aligned} \tag{4.41}$$

where the last line of (4.41) follows from (4.34). Using (4.26) in the right-hand side of (4.41), we have

$$\begin{aligned}\mathrm{E}\left\{\delta\mathbf{R}_d\mathbf{\Psi}\delta\mathbf{R}_d\right\} &= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{k=1}^{N}\mathrm{E}\big\{\mathbf{y}(n)\mathbf{\Psi}\mathbf{y}(k)\big\}\\ &= \frac{1}{N}\mathrm{E}\big\{\mathbf{y}(n)\mathbf{\Psi}\mathbf{y}(n)\big\} + \frac{1}{N}\sum_{\substack{k=1\\k\neq n}}^{N}\mathrm{E}\big\{\mathbf{y}(n)\mathbf{\Psi}\mathbf{y}(k)\big\}\end{aligned} \tag{4.42}$$

where

$$\mathbf{y}(n) = \mathbf{x}(n)\mathbf{x}(n)^H - \mathbf{J}\mathbf{x}(n)^*\mathbf{x}(n)^T\mathbf{J}. \tag{4.43}$$

Assuming that the received samples corresponding to two different sampling times are independent from each other, the expression inside the expectation operation in the last line of (4.42) is equal to $\mathbf{R}_d\mathbf{\Psi}\mathbf{R}_d$, which, according to (4.34), is zero. Hence, we have

$$\begin{aligned}\mathrm{E}\left\{\delta\mathbf{R}_d\mathbf{\Psi}\delta\mathbf{R}_d\right\} &= \frac{1}{N}\mathrm{E}\big\{\mathbf{y}(n)\mathbf{\Psi}\mathbf{y}(n)\big\} \\ &= \frac{1}{N}\left\{\mathbb{S}(\mathbf{\Psi}) + \mathbf{J}\mathbb{S}^T(\mathbf{J}\mathbf{\Psi}^T\mathbf{J})\mathbf{J} - \mathbb{T}(\mathbf{\Psi}) - \mathbb{T}^H(\mathbf{\Psi})\right\}\end{aligned} \tag{4.44}$$

where

$$\mathbb{S}(\mathbf{\Psi}) = \mathrm{E}\big\{\mathbf{x}\mathbf{x}^H\mathbf{\Psi}\mathbf{x}\mathbf{x}^H\big\} \tag{4.45}$$

$$\mathbb{T}(\mathbf{\Psi}) = \mathrm{E}\big\{\mathbf{x}\mathbf{x}^H\mathbf{\Psi}\mathbf{J}\mathbf{x}^*\mathbf{x}^T\mathbf{J}\big\} \tag{4.46}$$

and for the sake of simplicity, the time index n in $\mathbf{x}(n)$ has been omitted. Inserting (2.6) in (4.45) and (4.46), and using (4.33) along with the fact that $\mathbf{b}$ and $\mathbf{v}$ are zero-mean independent random vectors, we have

$$\mathbb{S}(\mathbf{\Psi}) = \mathrm{E}\big\{\mathbf{W}\mathbf{b}\mathbf{v}^H\mathbf{\Psi}\mathbf{v}\mathbf{b}^H\mathbf{W}^H\big\} + \mathrm{E}\big\{\mathbf{v}\mathbf{v}^H\mathbf{\Psi}\mathbf{v}\mathbf{v}^H\big\} \tag{4.47}$$

$$\mathbb{T}(\mathbf{\Psi}) = \mathrm{E}\big\{\mathbf{W}\mathbf{b}\mathbf{v}^H\mathbf{\Psi}\mathbf{J}\mathbf{v}^*\mathbf{b}^T\mathbf{W}^T\mathbf{J}\big\} + \mathrm{E}\big\{\mathbf{v}\mathbf{v}^H\mathbf{\Psi}\mathbf{J}\mathbf{v}^*\mathbf{v}^T\mathbf{J}\big\}. \tag{4.48}$$

Computing the first term of $\mathbb{S}(\mathbf{\Psi})$, we have

$$\begin{aligned}\mathrm{E}\big\{\mathbf{W}\mathbf{b}\mathbf{v}^H\mathbf{\Psi}\mathbf{v}\mathbf{b}^H\mathbf{W}^H\big\} &= \mathrm{E}\big\{\mathbf{v}^H\mathbf{\Psi}\mathbf{v}\big\}\mathrm{E}\big\{\mathbf{W}\mathbf{b}\mathbf{b}^H\mathbf{W}^H\big\} \\ &= \mathrm{tr}\big(\mathrm{E}\big\{\mathbf{v}\mathbf{v}^H\mathbf{\Psi}\big\}\big)\mathrm{E}\big\{\mathbf{W}\mathbf{b}\mathbf{b}^H\mathbf{W}^H\big\} \\ &= \mathrm{tr}(\mathbf{\Sigma}_v\mathbf{\Psi})\mathbf{W}\mathbf{W}^H.\end{aligned} \tag{4.49}$$

As the original RF version of noise is a wide-sense stationary random process, $\mathbf{v}$ is a circular random vector [56], and, therefore, the first term in the right-hand side of (4.48) is equal to zero. Using the so-obtained values of $\mathbb{S}(\mathbf{\Psi})$ and $\mathbb{T}(\mathbf{\Psi})$ in (4.44), after straightforward manipulations it can be shown that

$$\mathrm{E}\left\{\delta\mathbf{R}_d\mathbf{\Psi}\delta\mathbf{R}_d\right\} = \frac{1}{N}\big(\mathrm{tr}\left(\mathbf{\Sigma}_v\mathbf{\Psi}\right)(\mathbf{W}\mathbf{W}^H + \mathbf{J}\mathbf{W}^*\mathbf{W}^T\mathbf{J}) + \mathbf{\Theta}\big) \tag{4.50}$$

where $\mathbf{\Theta}$ is defined in (4.32). Substituting (4.50) in (4.40), equation (4.30) directly follows. This completes our proof. □

It should be noted that, as the noise power goes to zero, both $\mathbf{\Sigma}_v$ and $\mathbf{\Theta}$ converge to zero matrices. In such a case, it can be observed from (4.30) that even for the finite number of data samples, the approximate MSE of the channel vector estimate tends to zero. It is also noteworthy that, for any arbitrary wide-sense stationary noise, the MSE in (4.30) converges to zero with the rate $\mathcal{O}(1/N)$.

To further simplify the analysis of properties of the MSE expression (4.30), let us assume for the remainder of this section that $\mathbf{v}(n)$ has a circular Gaussian distribution, and, therefore, the fourth-order moments of $\mathbf{v}(n)$ are representable in terms of the second-order moments [4], [30]. Note that circular Gaussian noise has been frequently considered in the literature on interference rejection for CDMA systems [11], [20], [21], [67], [105]. Before proceeding to obtain the MSE of the channel estimation error for a circular Gaussian interference, we need the following lemma.

Lemma 2: Assume that $\mathbf{y}$ is a zero-mean circular Gaussian random vector with the correlation matrix $\mathbf{G}$, and $\mathbf{\Xi}$ is a square matrix of a conformable dimension. Then

$$\mathrm{E}\left\{\mathbf{y}\mathbf{y}^H\mathbf{\Xi}\mathbf{y}\mathbf{y}^H\right\} = \mathbf{G}\mathbf{\Xi}\mathbf{G} + \mathrm{tr}\left(\mathbf{G}\mathbf{\Xi}\right)\mathbf{G} \tag{4.51}$$

$$\mathrm{E}\left\{\mathbf{y}\mathbf{y}^H\mathbf{\Xi}\mathbf{y}^*\mathbf{y}^T\right\} = \mathbf{G}\mathbf{\Xi}\mathbf{G}^T + \mathbf{G}\mathbf{\Xi}^T\mathbf{G}^T. \tag{4.52}$$

Proof: Assume that $\mathbf{y}$ is a vector with the length m. To prove (4.51), we note that

$$\begin{aligned}
[\mathbf{y}\mathbf{y}^H\mathbf{\Xi}\mathbf{y}\mathbf{y}^H]_{jk} &= \sum_{l=1}^{m}[\mathbf{y}\mathbf{y}^H\mathbf{\Xi}]_{jl}[\mathbf{y}\mathbf{y}^H]_{lk} \\
&= \sum_{l=1}^{m}\sum_{g=1}^{m}[\mathbf{y}\mathbf{y}^H]_{jg}[\mathbf{\Xi}]_{gl}[\mathbf{y}\mathbf{y}^H]_{lk} \\
&= \sum_{l=1}^{m}\sum_{g=1}^{m}[\mathbf{\Xi}]_{gl}[\mathbf{y}]_j[\mathbf{y}]_g^*[\mathbf{y}]_l[\mathbf{y}]_k^*.
\end{aligned} \tag{4.53}$$

Hence, we have

$$\begin{aligned}\mathrm{E}\big\{[\mathbf{y}\mathbf{y}^H\boldsymbol{\Xi}\mathbf{y}\mathbf{y}^H]_{jk}\big\} &= \sum_{l=1}^{m}\sum_{g=1}^{m}[\boldsymbol{\Xi}]_{gl}\mathrm{E}\big\{[\mathbf{y}]_j[\mathbf{y}]_g^*[\mathbf{y}]_l[\mathbf{y}]_k^*\big\} \\ &= \sum_{l=1}^{m}\sum_{g=1}^{m}[\boldsymbol{\Xi}]_{gl}\big(\mathrm{E}\big\{[\mathbf{y}]_j[\mathbf{y}]_g^*\big\}\,\mathrm{E}\{[\mathbf{y}]_l[\mathbf{y}]_k^*\big\} \\ &\quad +\mathrm{E}\{[\mathbf{y}]_j[\mathbf{y}]_k^*\}\,\mathrm{E}\big\{[\mathbf{y}]_l[\mathbf{y}]_g^*\big\}\big) \qquad (4.54)\end{aligned}$$

where the second line of (4.54) holds due to the fact that $\mathbf{y}$ is a circular Gaussian random vector (for example, see [30], p. 68). As $\mathbf{G}$ is the covariance matrix of $\mathbf{y}$, we have

$$\begin{aligned}\mathrm{E}\left\{[\mathbf{y}\mathbf{y}^H\boldsymbol{\Xi}\mathbf{y}\mathbf{y}^H]_{jk}\right\} &= \sum_{l=1}^{m}\sum_{g=1}^{m}[\boldsymbol{\Xi}]_{gl}[\mathbf{G}]_{jg}[\mathbf{G}]_{lk} + \\ &\quad [\boldsymbol{\Xi}]_{gl}[\mathbf{G}]_{jk}[\mathbf{G}]_{lg} \\ &= \sum_{l=1}^{m}[\mathbf{G}\boldsymbol{\Xi}]_{jl}[\mathbf{G}]_{lk} + [\mathbf{G}\boldsymbol{\Xi}]_{ll}[\mathbf{G}]_{jk} \\ &= [\mathbf{G}\boldsymbol{\Xi}\mathbf{G}]_{jk} + \mathrm{tr}(\mathbf{G}\boldsymbol{\Xi})[\mathbf{G}]_{jk}\end{aligned}$$

which directly proves (4.51). Using derivations similar to (4.53) and (4.54), we obtain

$$\begin{aligned}\mathrm{E}\left\{[\mathbf{y}\mathbf{y}^H\boldsymbol{\Xi}\mathbf{y}^*\mathbf{y}^T]_{jk}\right\} &= \sum_{l=1}^{m}\sum_{g=1}^{m}[\boldsymbol{\Xi}]_{gl}[\mathbf{G}]_{jg}[\mathbf{G}]_{kl} + \\ &\quad [\boldsymbol{\Xi}]_{gl}[\mathbf{G}]_{jl}[\mathbf{G}]_{kg} \\ &= \sum_{l=1}^{m}[\mathbf{G}\boldsymbol{\Xi}]_{jl}[\mathbf{G}]_{kl} + [\mathbf{G}\boldsymbol{\Xi}]_{kl}[\mathbf{G}]_{jl} \\ &= \sum_{l=1}^{m}[\mathbf{G}\boldsymbol{\Xi}]_{jl}[\mathbf{G}^T]_{lk} + [\mathbf{G}\boldsymbol{\Xi}]_{kl}[\mathbf{G}^T]_{lj} \\ &= [\mathbf{G}\boldsymbol{\Xi}\mathbf{G}^T]_{jk} + [\mathbf{G}\boldsymbol{\Xi}\mathbf{G}^T]_{kj} \\ &= [\mathbf{G}\boldsymbol{\Xi}\mathbf{G}^T]_{jk} + [\mathbf{G}\boldsymbol{\Xi}^T\mathbf{G}^T]_{jk}.\end{aligned}$$

This proves (4.52). □

In the following theorem, we will use Lemma 2 along with (4.30) to obtain an approximate expression for the MSE of the channel vector estimate in the presence of unknown correlated Gaussian noise. A high SNR approximation of the so-obtained MSE expression will also be derived.

Theorem 2: Assume that $\mathbf{v}(n)$ is a circular Gaussian random vector and $\mathbf{h}_1$ is estimated using (4.29). Then, the first-order perturbation theory-based approximation of the MSE of the estimation error is given by

$$\begin{aligned}\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\mathrm{tr}\left(\mathbf{\Sigma}_v\mathbf{\Psi}\right)}{N}\mathbf{w}_1^H\mathbf{R}_d^\dagger\left(\mathbf{R}+\mathbf{J}\mathbf{R}^T\mathbf{J}\right)\mathbf{R}_d^\dagger\mathbf{w}_1 \\ -\frac{2}{N}\mathbf{w}_1^H\mathbf{R}_d^\dagger\mathbf{J}(\mathbf{\Sigma}_v\mathbf{\Psi}\mathbf{\Sigma}_v)^T\mathbf{J}\mathbf{R}_d^\dagger\mathbf{w}_1\,. \end{aligned} \tag{4.55}$$

Moreover, if (2.22) along with the following two conditions hold

$$\mathbf{w}_i^H\mathbf{J}\mathbf{w}_j^* = 0 \tag{4.56}$$

$$\lambda_{\max}(\mathbf{\Sigma}_v) \ll \frac{1}{2}A_1^2\|\mathbf{w}_1\|^2 \tag{4.57}$$

then (4.55) can be simplified to

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\mathrm{tr}\left(\mathbf{\Sigma}_v\mathbf{\Psi}\right)}{NA_1^2}. \tag{4.58}$$

Proof: Using (4.51) and (4.52) in (4.32), it can be readily shown that for any circular Gaussian interference vector $\mathbf{v}$, we have

$$\mathbf{\Theta} = 2\,\mathrm{tr}\left(\mathbf{\Sigma}_v\mathbf{\Psi}\right)\mathbf{\Sigma}_v - 2\mathbf{J}(\mathbf{\Sigma}_v\mathbf{\Psi}\mathbf{\Sigma}_v)^T\mathbf{J}. \tag{4.59}$$

Substituting (4.59) into (4.30) and using (4.1), the approximation (4.55) follows. To prove (4.58), note that if (2.23) and (4.56) hold true, then, according to (4.16), $[\mathbf{W}\ \ \mathbf{J}\mathbf{W}^*]$ is an orthogonal matrix whose columns span $\mathtt{range}(\mathbf{R}_d)$, and hence, the columns of $\mathbf{U}_s$ are the normalized versions of the corresponding columns of $[\mathbf{W}\ \ \mathbf{J}\mathbf{W}^*]$. Therefore, we have

$$\mathbf{U}_s = \left[\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\ \ \cdots\ \ \frac{\mathbf{w}_K}{\|\mathbf{w}_K\|}\ \ \frac{\mathbf{J}\mathbf{w}_1^*}{\|\mathbf{w}_1\|}\ \ \cdots\ \ \frac{\mathbf{J}\mathbf{w}_K^*}{\|\mathbf{w}_K\|}\right] \tag{4.60}$$

$$\mathbf{\Lambda}_s = \mathrm{diag}\left\{A_1^2\|\mathbf{w}_1\|^2,\ldots,A_K^2\|\mathbf{w}_K\|^2,-A_1^2\|\mathbf{w}_1\|^2,\ldots,-A_K^2\|\mathbf{w}_K\|^2\right\}. \tag{4.61}$$

Using (4.60) and (4.61) in (4.38) yields

$$\mathbf{w}_1^H\mathbf{R}_d^\dagger = \frac{\mathbf{w}_1^H}{A_1^2\|\mathbf{w}_1\|^2}. \tag{4.62}$$

Equations (2.23) and (4.56) can be used along with (4.62) to obtain the following equalities:

$$\begin{aligned}\mathbf{w}_1^H\mathbf{R}_d^\dagger\mathbf{R}\mathbf{R}_d^\dagger\mathbf{w}_1 &= \frac{\mathbf{w}_1^H\mathbf{R}\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\\ &= \frac{\mathbf{w}_1^H\mathbf{W}\mathbf{W}^H\mathbf{w}_1+\mathbf{w}_1^H\boldsymbol{\Sigma}_v\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\\ &= \frac{A_1^2\|\mathbf{w}_1\|^4+\mathbf{w}_1^H\boldsymbol{\Sigma}_v\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\\ &= \frac{1}{A_1^2}\left(1+\frac{\mathbf{w}_1^H\boldsymbol{\Sigma}_v\mathbf{w}_1}{A_1^2\|\mathbf{w}_1\|^4}\right)\end{aligned} \tag{4.63}$$

$$\begin{aligned}\mathbf{w}_1^H\mathbf{R}_d^\dagger\mathbf{J}\mathbf{R}^T\mathbf{J}\mathbf{R}_d^\dagger\mathbf{w}_1 &= \frac{\mathbf{w}_1^H\mathbf{J}\mathbf{R}^T\mathbf{J}\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\\ &= \frac{\mathbf{w}_1^H\mathbf{J}\left(\mathbf{W}^*\mathbf{W}^T+\boldsymbol{\Sigma}_v^*\right)\mathbf{J}\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\\ &= \frac{\mathbf{w}_1^H\boldsymbol{\Sigma}_v\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\end{aligned} \tag{4.64}$$

$$\mathbf{w}_1^H\mathbf{R}_d^\dagger\mathbf{J}(\boldsymbol{\Sigma}_v\boldsymbol{\Psi}\boldsymbol{\Sigma}_v)^T\mathbf{J}\mathbf{R}_d^\dagger\mathbf{w}_1 = \frac{\mathbf{w}_1^H\mathbf{J}\left(\boldsymbol{\Sigma}_v\boldsymbol{\Psi}\boldsymbol{\Sigma}_v\right)^T\mathbf{J}\mathbf{w}_1}{A_1^4\|\mathbf{w}_1\|^4}\,. \tag{4.65}$$

Using (4.63)-(4.65) in (4.55), it follows that

$$\mathrm{E}\left\{\|\delta\mathbf{h}_1\|^2\right\} \approx \frac{\mathrm{tr}(\boldsymbol{\Sigma}_v\boldsymbol{\Psi})}{NA_1^2}\left(1+\frac{2\xi}{A_1^2\|\mathbf{w}_1\|^4}\right) \tag{4.66}$$

where

$$\xi = \mathbf{w}_1^H\boldsymbol{\Sigma}_v\mathbf{w}_1 - \frac{\mathbf{w}_1^H\mathbf{J}\left(\boldsymbol{\Sigma}_v\boldsymbol{\Psi}\boldsymbol{\Sigma}_v\right)^T\mathbf{J}\mathbf{w}_1}{\mathrm{tr}(\boldsymbol{\Sigma}_v\boldsymbol{\Psi})}\,. \tag{4.67}$$

For any vector $\mathbf{w}_1$ and any positive (semi-)definite matrix $\boldsymbol{\Sigma}_v$, we have

$$0 \le \mathbf{w}_1^H\boldsymbol{\Sigma}_v\mathbf{w}_1 \le \|\mathbf{w}_1\|^2\lambda_{\max}(\boldsymbol{\Sigma}_v). \tag{4.68}$$

Moreover,

$$\begin{aligned} 0 \leq \frac{\mathbf{w}_1^H \mathbf{J} \left(\boldsymbol{\Sigma}_v \boldsymbol{\Psi} \boldsymbol{\Sigma}_v\right)^T \mathbf{J} \mathbf{w}_1}{\mathrm{tr}(\boldsymbol{\Sigma}_v \boldsymbol{\Psi})} &\leq \frac{\|\mathbf{w}_1\|^2 \lambda_{\max}\left(\mathbf{J}\left(\boldsymbol{\Sigma}_v \boldsymbol{\Psi} \boldsymbol{\Sigma}_v\right)^T \mathbf{J}\right)}{\mathrm{tr}(\boldsymbol{\Sigma}_v \boldsymbol{\Psi})} \\ &= \frac{\|\mathbf{w}_1\|^2 \lambda_{\max}\left(\mathbf{J}^2 \left(\boldsymbol{\Sigma}_v \boldsymbol{\Psi} \boldsymbol{\Sigma}_v\right)^T\right)}{\mathrm{tr}(\boldsymbol{\Sigma}_v \boldsymbol{\Psi})} \\ &= \frac{\|\mathbf{w}_1\|^2 \lambda_{\max}\left(\boldsymbol{\Sigma}_v \boldsymbol{\Psi} \boldsymbol{\Sigma}_v\right)}{\mathrm{tr}(\boldsymbol{\Sigma}_v \boldsymbol{\Psi})} \\ &\leq \frac{\|\mathbf{w}_1\|^2 \lambda_{\max}\left(\boldsymbol{\Sigma}_v \boldsymbol{\Psi}\right) \lambda_{\max}\left(\boldsymbol{\Sigma}_v\right)}{\mathrm{tr}(\boldsymbol{\Sigma}_v \boldsymbol{\Psi})} \\ &\leq \|\mathbf{w}_1\|^2 \lambda_{\max}(\boldsymbol{\Sigma}_v) \end{aligned} \tag{4.69}$$

where the third line of (4.69) follows from the facts that $\mathbf{J}^2 = \mathbf{I}$ and $\left(\boldsymbol{\Sigma}_v \boldsymbol{\Psi} \boldsymbol{\Sigma}_v\right)^T$ is Hermitian. Equations (4.68) and (4.69) yield

$$|\xi| \leq \|\mathbf{w}_1\|^2 \lambda_{\max}(\boldsymbol{\Sigma}_v). \tag{4.70}$$

Assuming that (4.57) holds, equation (4.58) immediately follows from inserting (4.70) into (4.66). This completes our proof. □

Note that (4.58) is a simple approximation of (4.55) in the high SNR regime that explicitly clarifies the MSE of the estimated channel vector in terms of the number of data samples, the received amplitude of the user-of-interest, and $\mathrm{tr}(\boldsymbol{\Sigma}_v \boldsymbol{\Psi})$. Note that the latter quantity can be viewed as a weighted average interference power where the weighting factor $\boldsymbol{\Psi}$ depends on the matrix $\boldsymbol{\mathcal{C}}_1$ and the principal angles between $\mathtt{range}(\boldsymbol{\mathcal{C}}_1)$ and $\mathtt{range}(\mathbf{U}_n)$ [47].

Assuming that the average noise power is equal to e_o, that is,

$$\mathrm{tr}\big(\boldsymbol{\Sigma}_v\big) = e_o, \tag{4.71}$$

and $\mathbf{h}_1$ is the unique nontrivial solution to (4.21), one can also obtain an upper-bound for the MSE in (4.58) as follows. As (4.21) has a unique nontrivial solution, it directly follows that

$$\mathtt{rank}(\mathbf{T}) = L - 1. \tag{4.72}$$

Let us denote the positive singular values of $\mathbf{T}$ as $\xi_1 \geq \xi_2 \geq \ldots \geq \xi_{L-1} > 0$. Note that the positive eigenvalues of $\mathbf{\Psi}$ and those of ${\mathbf{T}^\dagger}^H\mathbf{T}^\dagger$ are equal to $\xi_{L-1}^{-2} \geq \xi_{L-2}^{-2} \geq \cdots \geq \xi_1^{-2}$. Since $\mathbf{\Sigma}_v$ and $\mathbf{\Psi}$ are positive (semi-)definite matrices, we have [34]

$$\begin{aligned}\frac{\mathrm{tr}\,(\mathbf{\Sigma}_v\mathbf{\Psi})}{NA_1^2} &\leq \frac{\mathrm{tr}\big(\mathbf{\Sigma}_v\big)\lambda_{\max}(\mathbf{\Psi})}{NA_1^2} \\ &= \frac{e_o\lambda_{\max}({\mathbf{T}^\dagger}^H\mathbf{T}^\dagger)}{NA_1^2} \\ &= \frac{e_o}{NA_1^2\xi_{L-1}^2}. \end{aligned} \tag{4.73}$$

It should be noted that if the largest eigenvalue of $\mathbf{\Psi}$ is unique, i.e., $\xi_{L-1}^{-2} > \xi_{L-2}^{-2}$, then the first line of (4.73) holds with equality if and only if

$$\mathbf{\Sigma}_v = e_o\mathbf{s}\mathbf{s}^H \tag{4.74}$$

where $\mathbf{s}$ is the eigenvector of $\mathbf{\Psi}$ associated with ξ_{L-1}^{-2}. It follows from (4.73) that the MSE of the estimated channel vector can become very large if ξ_{L-1} goes to zero. This is an expected result since if $\xi_{L-1} = 0$, then $\mathtt{rank}(\mathbf{T}) = L-2$, and, therefore, (4.21) has a nontrivial solution other than $\mathbf{h}_1$.

4.4 Temporally oversampled version of the proposed algorithm

Condition (4.23) restricts both the maximum admissible number of active users and the channel length. However, one can resort to the temporal oversampling technique to facilitate identification of longer channels in more heavily loaded environments. Similar to Section 2.3.1, let us sample the received signal (1.1) with the rate κ/T_c in the interval corresponding to the nth transmitted symbol where the oversampling factor $\kappa \geq 1$ is an integer. The covariance matrix of the ISI-free part of the oversampled data vector is given by

$$\mathbf{R}_\kappa = \mathrm{E}\{\mathbf{x}_\kappa(n)\mathbf{x}_\kappa(n)^H\} = \mathbf{W}_\kappa\mathbf{W}_\kappa^H + \mathbf{\Sigma}_{v\kappa} \tag{4.75}$$

where

$$\mathbf{\Sigma}_{v\kappa} = \mathrm{E}\{\mathbf{v}_\kappa(n)\mathbf{v}_\kappa(n)^H\} \tag{4.76}$$

and $\mathbf{x}_\kappa(n)$, $\mathbf{v}_\kappa(n)$, and $\mathbf{W}_\kappa$ are given by (2.30), (2.32), and (2.41), respectively. Note that $\mathbf{R}_\kappa$ is a $\kappa(L_c - L + 1) \times \kappa(L_c - L + 1)$ matrix and

$$\mathtt{dim}\{\mathtt{range}(\mathbf{W}_\kappa)\} = K. \tag{4.77}$$

Introducing

$$\mathbf{R}_{\kappa d} = \mathbf{R}_\kappa - \mathbf{J}\mathbf{R}_\kappa^*\mathbf{J} \tag{4.78}$$

and using (4.75) and the centro-Hermitian property of $\mathbf{\Sigma}_{v\kappa}$, we have

$$\mathbf{R}_{\kappa d} = \begin{bmatrix} \mathbf{W}_\kappa & \mathbf{J}\mathbf{W}_\kappa^* \end{bmatrix} \begin{bmatrix} \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_K \end{bmatrix} \begin{bmatrix} \mathbf{W}_\kappa & \mathbf{J}\mathbf{W}_\kappa^* \end{bmatrix}^H. \tag{4.79}$$

The matrix $\mathbf{R}_{\kappa d}$ can be eigendecomposed as

$$\mathbf{R}_{\kappa d} = \begin{bmatrix} \mathbf{U}_{\kappa s} & \mathbf{U}_{\kappa n} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{\kappa s} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\kappa s}^H \\ \mathbf{U}_{\kappa n}^H \end{bmatrix} \tag{4.80}$$

where $\mathbf{U}_{\kappa n}$ is a $\kappa(L_c - L + 1) \times (\kappa(L_c - L + 1) - 2K)$ matrix which contains the eigenvectors associated with the zero eigenvalue of $\mathbf{R}_{\kappa d}$. Similar to our discussion in Section 4.2, it can be easily shown that $\mathbf{h}_{1\kappa}$, the oversampled version of the channel vector introduced in (2.38), is a nontrivial solution to

$$\mathbf{T}_\kappa \mathbf{h} = \mathbf{0} \tag{4.81}$$

where

$$\mathbf{T}_\kappa = \mathbf{U}_{n\kappa}^H \boldsymbol{\mathcal{C}}_{1\kappa} \tag{4.82}$$

is a $(\kappa(L_c - L + 1) - 2K) \times \kappa L$ matrix and $\boldsymbol{\mathcal{C}}_{1\kappa}$ is given by (2.37). To have a unique nontrivial solution for (4.81), it is required that

$$\kappa(L_c + 1) \geq 2\kappa L + 2K. \tag{4.83}$$

Comparing (4.83) with (4.23), it follows that using the oversampling factor of κ, the maximum admissible number of users can be increased by κ times. It can also be observed from (4.83) that for a fixed number of active users, the oversampling technique considerably increases the maximum admissible channel length.

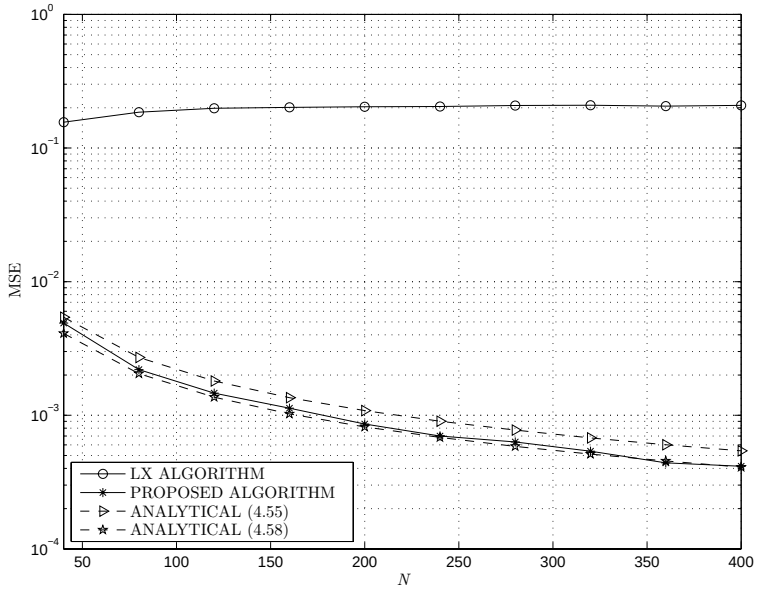

Figure 4.1: MSEs of the estimated channel versus the number of data samples N for the first noise model.

4.5 Simulations

Computer simulations have been carried out to evaluate the performance of the proposed algorithm and validate the obtained theoretical results. In all numerical examples, $L_c = 40$ and the spreading sequences associated with each user have been randomly drawn from the binary set of $\{-1, +1\}$ and then fixed throughout all examples. Similarly, the entries of the channel vectors of the length $L = 4$ have been randomly and independently drawn from a zero-mean complex Gaussian process and then have been normalized so that $\|\mathbf{h}_k\| = 1$ ($k = 1, \ldots, K$) and fixed throughout all examples. In all but the last example, the transmitted symbols have been drawn from the QPSK constellation. Three different models have been used to simulate the interference vector:

- **Model 1:** The noise vector $\mathbf{v}(n)$ is a circular Gaussian random vector such that the (l, k)-th entry of the correlation matrix $\mathbf{\Sigma}_v$ is [42]

$$[\mathbf{\Sigma}_v]_{lk} = 0.98^{|l-k|}. \tag{4.84}$$

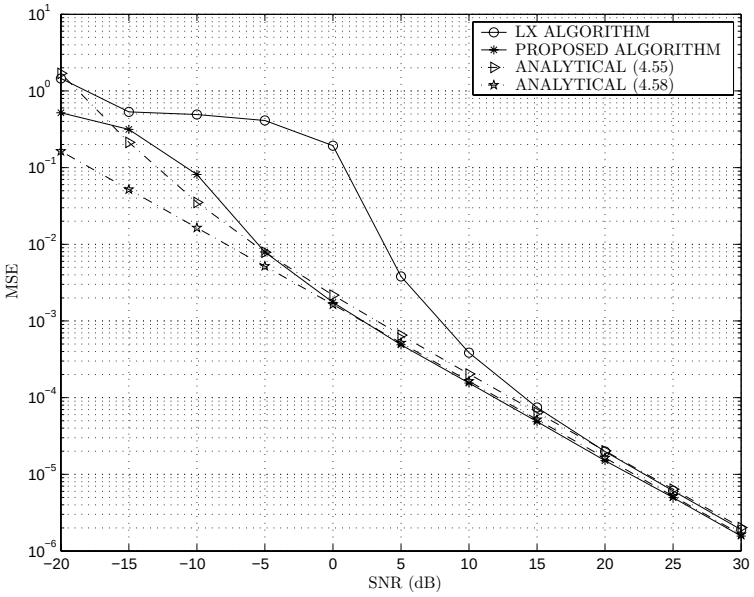

Figure 4.2: MSEs of the estimated channel versus SNR for the first noise model.

- **Model 2:** The correlated noise is a second-order Gaussian autoregressive process [11], [21] with the poles at $p_1 = 0.95$ and $p_2 = 0.97$, that is,

$$[\mathbf{v}(n)]_k - 1.92[\mathbf{v}(n)]_{k-1} + 0.9215[\mathbf{v}(n)]_{k-2} = [\mathbf{e}(n)]_k \tag{4.85}$$

  where $\mathbf{e}(n)$ is the vector of white complex Gaussian noise.

- **Model 3:** The correlated noise is tonal (harmonic) with

$$[\mathbf{v}(n)]_k = e^{j(\Omega k + \theta(n))} \tag{4.86}$$

  where $\Omega = \pi/6$ describes the normalized noise frequency offset from the carrier frequency, and $\theta(n)$ is a random phase uniformly distributed in the interval $[0, 2\pi]$ [62], [39].

In Figs. 4.1 through 4.3, the first noise model has been used, while in Figs. 4.4 and 4.5, the second noise model has been applied. In Fig. 4.6, the third noise model is employed. Throughout our simulations, we assume that all the users have identical average powers and $K = 5$ is taken in all the figures but Fig. 4.3. Each point of the simulation curves is

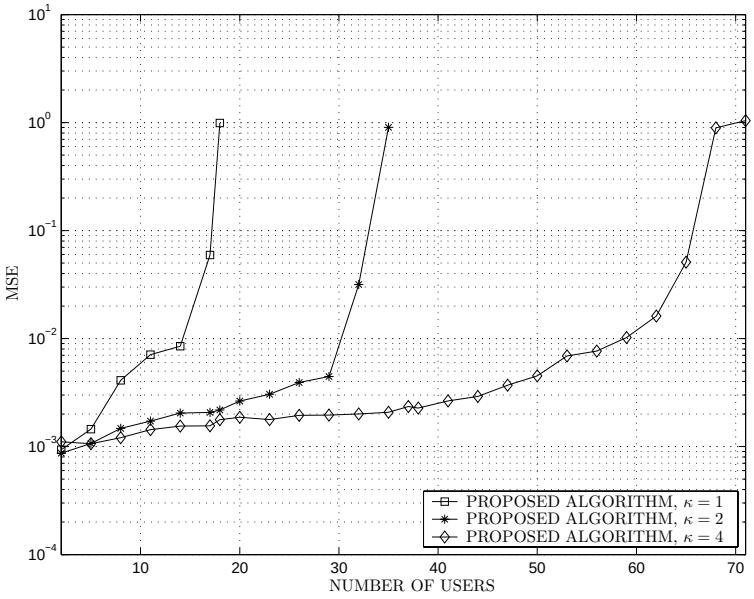

Figure 4.3: MSEs of the estimated channel versus number of users for the first interference model.

the result of averaging over 1000 Monte-Carlo realizations of the noise and the transmitted data sequences.

Fig. 4.1 shows the experimental MSE of the proposed algorithm as well as the theoretical MSEs (4.55) and (4.58) versus N for SNR = 0 dB. For the sake of comparison, the MSE curve of the conventional LX algorithm [47] is also drawn. It can be observed from Fig. 4.1 that the analytical MSE curves obtained from (4.55) and (4.58) follow the experimental MSE curve with quite a good accuracy. As predicted in Section 4.3, these curves converge to zero with the rate $1/N$. From Fig. 4.1, it also follows that the MSE of the estimated channel using the LX algorithm is constantly high and does not converge to zero. Note that the LX algorithm is based on the mismatched assumption that the signal subspace is spanned by the eigenvectors associated with the K largest eigenvalues of $\mathbf{R}$. Hence, even if N becomes arbitrarily large and $\hat{\mathbf{R}}$ converges to $\mathbf{R}$, the LX algorithm does not offer better estimation performance.

Fig. 4.2 shows the experimental and analytical MSEs versus SNR for $N = 100$. A substantial performance improvement can be observed from this figure with respect to the

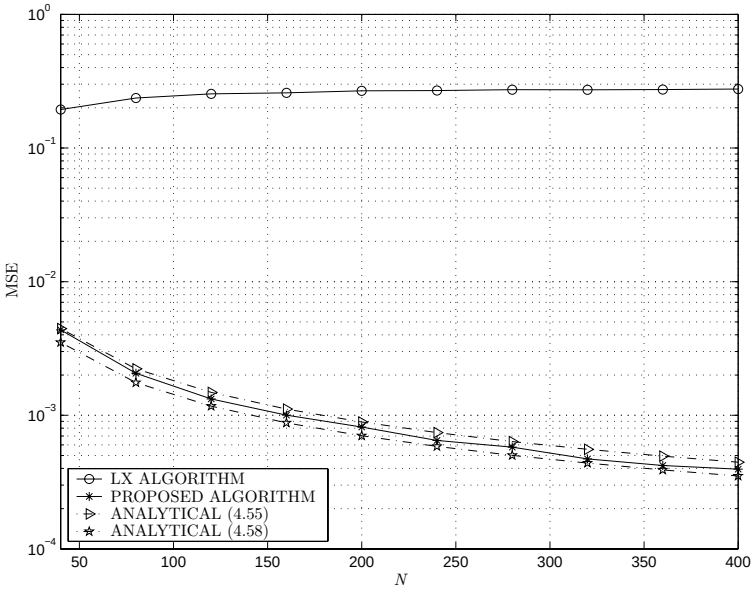

Figure 4.4: MSEs of the estimated channel versus number of data samples N for the second interference model.

LX algorithm. Note that the effect of noise is negligible in high SNRs where, as Fig. 4.2 demonstrates, the conventional LX algorithm can also be used to obtain a reliable channel vector estimate. Note also that the MSE expressions (4.55) and (4.58) are derived using the first-order perturbation theory, and, hence, their validity requires a small perturbation assumption. Therefore, as shown in Fig. 4.2, equations (4.55) and (4.58) cannot accurately predict the MSE values at very low SNRs, where the MSE of the channel vector estimate is quite large.

Fig. 4.3 displays the MSE of the proposed technique versus K for SNR = 0 dB and $N = 100$ in the case of oversampling. Three different oversampling factors of $\kappa = 1$ (no oversampling), $\kappa = 2$, and $\kappa = 4$ are considered. If no oversampling is used, then, as follows from (4.23), the proposed algorithm fails to estimate the channel vector for $K > 16$. This theoretical result is validated by Fig. 4.3. At the same time, it can be observed from Fig. 4.3 that using the oversampled version of the proposed algorithm with the oversampling factors of $\kappa = 2$ and $\kappa = 4$, the channel vector can be reliably estimated in the presence of up to $K = 30$ and $K = 60$ users, respectively.

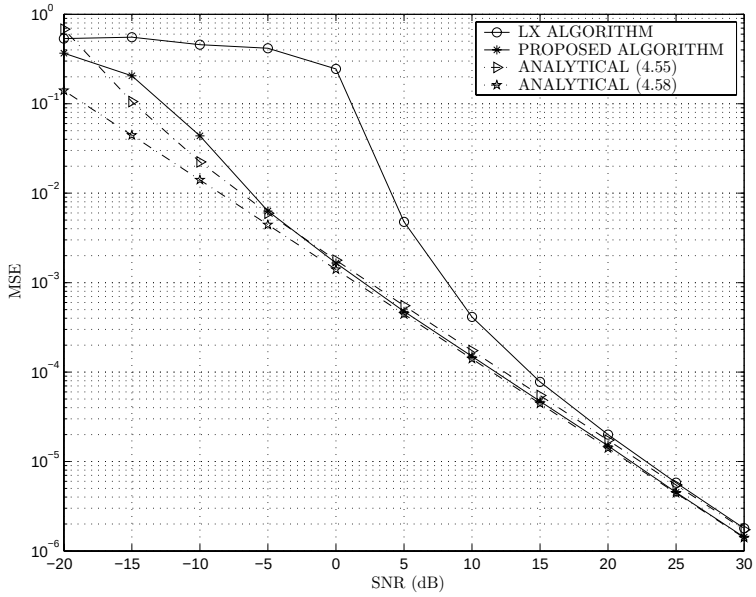

Figure 4.5: MSEs of the estimated channel versus SNR for the second interference model.

Figs. 4.4 and 4.5 show the MSE curves versus N and SNR, respectively, where (4.85) is used to model interference. Except the interference model, the simulation setup in Figs. 4.4 and 4.5 is identical to that of Figs. 4.1 and 4.2, respectively. Figs. 4.4 and 4.5 demonstrate that the proposed technique outperforms the conventional LX algorithm, and the analytical MSE curves coincide with the experimental MSE curves with a good precision.

Fig. 4.6 shows the MSE curves for the proposed algorithm and for the BP algorithm [11] versus SNR for both the cases of BPSK and QPSK transmitted symbols. The analytical MSE curve of the proposed algorithm is also shown in this figure. Note that, as the correlated noise is not Gaussian, the latter curve is obtained from (4.30). As it can be observed from Fig. 4.6, the proposed algorithm performs equally well for both the BPSK and QPSK transmission schemes. Moreover, both the experimental MSE curves are quite close to the analytical curve (4.30). It can also be seen from Fig. 4.6 that, although the performance of the BP algorithm for the BPSK transmission scheme is comparable to that of the proposed algorithm, the BP algorithm becomes completely unreliable in the case of QPSK symbols. This is, however, an expected result since the BP algorithm is exclusively designed for the

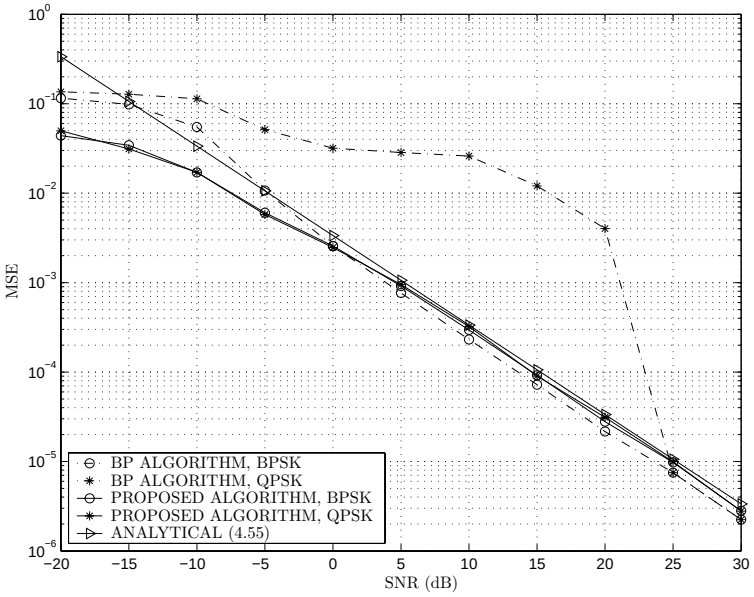

Figure 4.6: MSEs of the estimated channel versus SNR for the third interference model

BPSK symbol case [11].

4.6 Chapter summary

In this chapter, we have proposed a new blind subspace-based signature waveform estimation technique for DS-CDMA communication systems operating in the presence of unknown wide-sense stationary noise. Using the centro-Hermitian property of the noise covariance matrix along with the idea of covariance differencing [63], we derive a new algorithm for blind identification of user signatures. In contrast to the WP and BP algorithms, the proposed technique can be implemented using a single receiving antenna and is applicable to arbitrary constellations of transmitted symbols. Necessary and sufficient conditions for identifiability of the proposed technique have also been derived. Using the first-order perturbation theory, closed-form expressions for the MSE of the estimated channel vector have been obtained and the effects of different parameters on the performance of the proposed algorithm have been studied. It has been shown that the temporally oversampled version of our technique can be applied to overloaded communication systems with lengthy channels. Numerical examples

have verified the advantages of the proposed technique and validate its performance analysis.

Chapter 5

Robust blind MMSE multiuser detection

5.1 Introduction

Multiuser receivers require the knowledge of the data covariance matrix as well as the signature of the user-of-interest to perform the symbol detection [84]. However, as discussed in former chapters, neither of these values are exactly known at the receiver in practical scenarios. The data covariance matrix is usually estimated using the sampled received data while, as discussed before, the signature of the user-of-interest can be estimated using either training-based or blind techniques. Several practical hurdles such as the short available data length, variations in the channel impulse response during the observation period, non-stationarity of the received signal, and low SNR may result in considerable discrepancies between the actual and the estimated values of the user signature and the data covariance matrix. This, in turn, may cause a severe degradation of the detection performance.

One efficient approach to overcome this problem is to introduce robustness in the multiuser detection procedure [1], [16] [31], [70], [75], [86], [87]. A robust multiuser receiver which is based on the MOE approach has been presented in [31]. This receiver minimizes

the output power subject to linear and quadratic constraints, where the linear constraint guarantees distortionless response to the user-of-interest while the quadratic constraint limits the norm of the difference between the receiver coefficient vector and the presumed signature of the user-of-interest. In [75], another MOE-based multiuser receiver has been proposed where the receiver output power is minimized subject to a quadratic constraint (which limits the norm of the receiver coefficient vector) and a set of *ad hoc* linear constraints that are used to provide further robustness against signature mismatch and channel distortions.

It can be shown that the robust receivers proposed in [31] and [75] are both based on diagonal loading of the covariance matrix of the received data. However, the main shortcoming of this approach is that it is not clear how to obtain the optimal value of the diagonal loading factor. In [16], it has been shown that the robustness of the diagonal loading-based approach with standard ad-hoc choices of the diagonal loading factor may be insufficient. Motivated by this drawback of the diagonal loading technique, the authors of [16] have proposed another promising approach to robust blind multiuser detection that explicitly models an arbitrary (but norm-bounded) uncertainty in the signature of the user-of-interest and uses worst-case performance optimization to improve the robustness of the MOE receiver. This method is based on SOCP. Although several efficient convex optimization software tools are currently available, the SOCP-based method does not have any closed-form solution and is not suitable for on-line implementations. In fact, the receiver coefficients have to be recomputed completely whenever a new data sample is received.

In [70], two other robust MOE-based multiuser receivers are proposed which have closed-form solutions. These multiuser detection techniques are based on the optimization of a lower-bound of the worst-case performance and use the covariance matrix of the signature of the user-of-interest rather than the signature itself. These algorithms are shown to outperform the diagonal loading-based multiuser receiver, but their performances may be affected by the use of the lower bound on the worst-case performance rather than the worst-case performance itself.

In this chapter (see also [117], [118]), we use the MMSE multiuser detection approach

along with the idea of worst-case performance optimization to develop a new blind multiuser receiver which is robust against possible uncertainties in the MSE cost function. Our approach is based on the explicit modeling of uncertainties in both the signature of the user-of-interest and the data covariance matrix.[1] It is shown that this approach is equivalent to the diagonal loading-based multiuser receiver with the optimal choice of the diagonal loading factor that is obtained based on the known level of uncertainty in the signature of the user of interst. A computationally efficient algorithm is derived to compute the coefficient vector of the proposed robust multiuser receiver. In contrast to the SOCP-based algorithm [16], the latter technique is suitable for on-line implementation. We also show that the robust SOCP-based receiver of [16] is a special case of our receiver. Hence, the algorithm provided for the proposed receiver can also be used for implementing the receiver [16] in a computationally efficient way.

This chapter is organized as follows. A background on the MMSE multiuser receiver is presented in Section 5.2. In Section 5.3, we develop a robust formulation for the MMSE receiver and derive its computationally simple solution based on a Newton search technique. In the same section, the relationship between our multiuser receiver and the SOCP-based receiver of [16] is established. Section 5.4 contains simulation results. A summary of the contributions of this chapter is given in Section 5.5.

5.2 Background

5.2.1 Linear receivers

As the optimal ML detector may have a prohibitively high computational cost, the linear receivers are often employed as suboptimal but computationally attractive alternative solutions to the multiuser detection problem. The output of a linear multiuser receiver is given by [32], [49]

$$y(n) = \mathbf{f}^H \underline{\mathbf{x}}(n) \tag{5.1}$$

[1] Robust receivers may also be designed to counter other deteriorating effects such as that of the impulsive noise [57], [92], [119].

where $\mathbf{f}$ is an $L_c \times 1$ complex vector of the receiver coefficients. The receiver output $y(n)$ is used for symbol detection. For example, in BPSK systems, the symbol detection is made as [84]

$$\hat{b}_1(n) = \text{sgn}[\text{Re}(y(n))]. \tag{5.2}$$

5.2.2 MMSE receiver

In the MMSE approach [50], the receiver coefficient vector $\mathbf{f}$ is designed to minimize the MSE between the symbol of the user-of-interest and the receiver output. Therefore, the optimal coefficient vector $\mathbf{f}_{\text{MMSE}}$ is obtained as

$$\begin{aligned} \mathbf{f}_{\text{MMSE}} &= \arg\min_{\mathbf{f}} \text{E}\{|b_1(n) - \mathbf{f}^H \underline{\mathbf{x}}(n)|^2\} \\ &= \arg\min_{\mathbf{f}} \left\{ \mathbf{f}^H \underline{\mathbf{R}} \mathbf{f} - \underline{\mathbf{d}}^H \mathbf{f} - \mathbf{f}^H \underline{\mathbf{d}} \right\} \end{aligned} \tag{5.3}$$

where $\underline{\mathbf{R}} = \text{E}\{\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H\}$ and

$$\underline{\mathbf{d}} = \text{E}\{\underline{\mathbf{x}}(n) b_1^*(n)\} \tag{5.4}$$

is the correlation vector between the received data vector $\underline{\mathbf{x}}(n)$ and the conjugate of the symbol of the user-of-interest $b_1(n)$. The optimal coefficient vector is given by the classic Wiener formula:

$$\mathbf{f}_{\text{MMSE}} = \underline{\mathbf{R}}^{-1} \underline{\mathbf{d}}. \tag{5.5}$$

Note that inserting (1.7) into (5.4), we have that

$$\underline{\mathbf{d}} = A_1 \underline{\mathbf{w}}_1. \tag{5.6}$$

Substituting (5.6) into (5.5) and noting that multiplying $\mathbf{f}_{\text{MMSE}}$ in (5.5) by any positive constant does not affect the probability of error at the output of the symbol detector, we can write

$$\mathbf{f}_{\text{MMSE}} = \underline{\mathbf{R}}^{-1} \underline{\mathbf{w}}_1. \tag{5.7}$$

Note that, for the sake of simplicity and with a small abuse of notation, the same subscript as in the actual MMSE receiver (5.5) has been used for the rescaled MMSE receiver (5.7).

Note that although the receivers (5.5) and (5.7) have identical BER performances, the MSEs of (5.5) and (5.7) are different because (5.7) does not, in fact, minimize the MSE.

As discussed in the previous chapters, the exact knowledge of the signature of the user-of-interest $\underline{\mathbf{w}}_1$ is unavailable in the practical blind CDMA systems. In such a case, the blind version of the receiver (5.7) can be written as

$$\mathbf{f}_{\text{blind}} = \underline{\mathbf{R}}^{-1}\underline{\hat{\mathbf{w}}}_1 \tag{5.8}$$

where $\underline{\hat{\mathbf{w}}}_1$ is some blind estimate of $\underline{\mathbf{w}}_1$. Unfortunately, the receiver (5.8) is sensitive to the difference between $\underline{\mathbf{w}}_1$ and $\underline{\hat{\mathbf{w}}}_1$, i.e., the performance of (5.8) can degrade substantially in the presence of even a slight mismatch between these two vectors [16], [31], [75].

In practical situations, the exact data covariance matrix is also unavailable and is replaced by its sample estimate

$$\underline{\hat{\mathbf{R}}} = \frac{1}{N}\sum_{n=1}^{N}\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H. \tag{5.9}$$

Using $\underline{\hat{\mathbf{R}}}$ in lieu of $\underline{\mathbf{R}}$, the multiuser receivers (5.7) and (5.8) can be written as

$$\hat{\mathbf{f}}_{\text{MMSE}} = \underline{\hat{\mathbf{R}}}^{-1}\underline{\mathbf{w}}_1 \tag{5.10}$$

$$\mathbf{f}_{\text{blind}} = \underline{\hat{\mathbf{R}}}^{-1}\underline{\hat{\mathbf{w}}}_1. \tag{5.11}$$

In scenarios with a short data length, the performance of the multiuser receivers (5.10) and (5.11) can degrade substantially [75]. To provide robustness against short data length effects, it has been proposed in [31] and [75] to use the so-called diagonal loading technique whose essence is to replace $\underline{\hat{\mathbf{R}}}$ by $\underline{\hat{\mathbf{R}}}+\gamma\mathbf{I}$ where γ is the loading factor[2]. Using this approach, the diagonally loaded blind MMSE multiuser receiver can be written as [31], [75]

$$\mathbf{f}_{\text{dl}} = (\underline{\hat{\mathbf{R}}} + \gamma\mathbf{I})^{-1}\underline{\hat{\mathbf{w}}}_1. \tag{5.12}$$

Although the receiver (5.12) is known to (potentially) provide an improved robustness against short data length effects and signature mismatches, it is not clear how to choose the

[2]Note that the diagonal loading approach was originally proposed in application to adaptive beamforming [1]. See also [83] and references therein.

diagonal loading factor based on known levels of uncertainty in the data covariance matrix and the signature of the user-of-interest. As a result, the amount of diagonal loading is usually chosen in an ad hoc way [75] (typically, 10 to 15 dB above the noise power [86]) and this may negatively affect the resulting performance of the diagonal loading-based multiuser receivers.

5.3 Robust multiuser detection

5.3.1 Formulation

The MMSE multiuser receiver in (5.3) assumes that the correlation vector $\underline{\mathbf{d}}$ and the data covariance matrix $\underline{\mathbf{R}}$ are exactly known. In practice, these values are known with certain errors. In this section, we present a blind multiuser receiver which is robust against such errors in $\underline{\mathbf{d}}$ and $\underline{\mathbf{R}}$.

Let us consider the error between the actual signature of the user-of-interest $\underline{\mathbf{w}}_1$ and its estimated value $\underline{\hat{\mathbf{w}}}_1$,

$$\mathbf{e} = \underline{\mathbf{w}}_1 - \underline{\hat{\mathbf{w}}}_1. \tag{5.13}$$

We assume that the error vector $\mathbf{e}$ is norm-bounded by some known constant $\varepsilon > 0$:

$$\|\mathbf{e}\| \leq \varepsilon. \tag{5.14}$$

In practice, the problem of determining ε is much simpler than the problem of finding the error vector $\mathbf{e}$ itself. Note that, typically, some preliminary (coarse) knowledge about the wireless channel is available. This information can be obtained either from preliminary channel measurement campaigns or using numerical channel modelling. Moreover, as discussed in the previous chapters, it is usually known how the signature estimation error depends on the parameters such as SNR and the number of data samples used in the signature estimation procedure. The above information can be used to find a proper value for ε for the given type of the channel.

Inserting (5.13) into (5.6), we obtain that

$$\underline{\mathbf{d}} = \underline{\hat{\mathbf{d}}} + A_1\mathbf{e} \tag{5.15}$$

where

$$\hat{\underline{\mathbf{d}}} = A_1 \, \hat{\underline{\mathbf{w}}}_1 \tag{5.16}$$

is the presumed correlation vector.

In addition to signature errors, there is always a certain error between the true data covariance matrix $\underline{\mathbf{R}}$ and its sample estimate $\hat{\underline{\mathbf{R}}}$. Hence, we can write

$$\underline{\mathbf{R}} = \hat{\underline{\mathbf{R}}} + \mathbf{E} \tag{5.17}$$

where $\mathbf{E}$ is a Hermitian matrix which takes into account all errors that may be caused by short data length effects. We assume that $\mathbf{E}$ is bounded in its Frobenius norm by some known constant $\gamma > 0$, i.e.,

$$\|\mathbf{E}\|_F \leq \gamma. \tag{5.18}$$

In practice, a proper value of γ can be found using known results for the covariance of the matrix $\mathbf{E}$, see [99] and references therein.

To incorporate robustness against such norm-bounded errors $\mathbf{e}$ and $\mathbf{E}$, let us modify the MMSE optimization problem in (5.3) as

$$\min_{\mathbf{f}} \max_{\substack{\|\mathbf{E}\|_F \leq \gamma \\ \|\mathbf{e}\| \leq \varepsilon}} \Big\{ \mathbf{f}^H (\hat{\underline{\mathbf{R}}} + \mathbf{E}) \mathbf{f} - \mathbf{f}^H (\hat{\underline{\mathbf{d}}} + A_1 \mathbf{e}) - (\hat{\underline{\mathbf{d}}} + A_1 \mathbf{e})^H \mathbf{f} \Big\}. \tag{5.19}$$

The problem statement in (5.19) guarantees that the MSE cost function is minimized for the worst-case scenario which corresponds to the largest value of the MSE over all possible norm-bounded errors in the signature of the user-of-interest and the data covariance matrix. Therefore, the proposed design protects the MMSE receiver performance against worst-case errors and, correspondingly, it should improve its robustness.

Note that the objective function in (5.19) is written as the sum of two terms so that each of them depends on one of the optimization variables only (i.e., the first term depends on $\mathbf{E}$, whereas the second term depends on $\mathbf{e}$). Taking into account this fact, (5.19) can be

rewritten as

$$\min_{\mathbf{f}} \Big\{ \max_{\|\mathbf{E}\|_F \le \gamma} \{\mathbf{f}^H(\hat{\underline{\mathbf{R}}} + \mathbf{E})\mathbf{f}\} + \max_{\|\mathbf{e}\| \le \varepsilon} \{-\mathbf{f}^H(\hat{\underline{\mathbf{d}}} + A_1\mathbf{e}) - (\hat{\underline{\mathbf{d}}} + A_1\mathbf{e})^H\mathbf{f}\} \Big\}. \tag{5.20}$$

To simplify (5.20), the following two lemmas will be used.

Lemma 1: For any Hermitian $\mathbf{E}$ and $\hat{\underline{\mathbf{R}}}$ and any fixed $\mathbf{f}$,

$$\max_{\|\mathbf{E}\|_F \le \gamma} \mathbf{f}^H(\hat{\underline{\mathbf{R}}} + \mathbf{E})\mathbf{f} = \mathbf{f}^H(\hat{\underline{\mathbf{R}}} + \gamma\mathbf{I})\mathbf{f}. \tag{5.21}$$

Proof: See [71]. □

Lemma 2: For any fixed $\mathbf{f}$,

$$\max_{\|\mathbf{e}\| \le \varepsilon} \{-\mathbf{f}^H(\hat{\underline{\mathbf{d}}} + A_1\mathbf{e}) - (\hat{\underline{\mathbf{d}}} + A_1\mathbf{e})^H\mathbf{f}\} = -\mathbf{f}^H\hat{\underline{\mathbf{d}}} - \hat{\underline{\mathbf{d}}}^H\mathbf{f} + 2\varepsilon A_1 \|\mathbf{f}\|. \tag{5.22}$$

Proof: Let us write

$$\begin{aligned} &\max_{\|\mathbf{e}\| \le \varepsilon} \{-\mathbf{f}^H(\hat{\underline{\mathbf{d}}} + A_1\mathbf{e}) - (\hat{\underline{\mathbf{d}}} + A_1\mathbf{e})^H\mathbf{f}\} \\ &\quad = -\mathbf{f}^H\hat{\underline{\mathbf{d}}} - \hat{\underline{\mathbf{d}}}^H\mathbf{f} + A_1 \max_{\|\mathbf{e}\| \le \varepsilon} \{-\mathbf{f}^H\mathbf{e} - \mathbf{e}^H\mathbf{f}\} \\ &\quad = -\mathbf{f}^H\hat{\underline{\mathbf{d}}} - \hat{\underline{\mathbf{d}}}^H\mathbf{f} - A_1 \min_{\|\mathbf{e}\| \le \varepsilon} \{\mathbf{f}^H\mathbf{e} + \mathbf{e}^H\mathbf{f}\} \\ &\quad = -\mathbf{f}^H\hat{\underline{\mathbf{d}}} - \hat{\underline{\mathbf{d}}}^H\mathbf{f} - 2A_1 \min_{\|\mathbf{e}\| \le \varepsilon} \mathrm{Re}(\mathbf{f}^H\mathbf{e}) \,. \end{aligned} \tag{5.23}$$

Since for a complex number z we have $\mathrm{Re}(z) \ge -|z|$, the objective function of the minimization in (5.23) is lower bounded as

$$\begin{aligned} \mathrm{Re}(\mathbf{f}^H\mathbf{e}) &\ge -|\mathbf{f}^H\mathbf{e}\,| \\ &\ge -\|\mathbf{f}\|\|\mathbf{e}\| \\ &\ge -\varepsilon\|\mathbf{f}\| \end{aligned} \tag{5.24}$$

where the Cauchy-Schwartz inequality and the constraint $\|\mathbf{e}\| \le \varepsilon$ have been used. From (5.24), it follows that

$$\min_{\|\mathbf{e}\| \le \varepsilon} \mathrm{Re}(\mathbf{f}^H\mathbf{e}) \ge -\varepsilon\|\mathbf{f}\| \tag{5.25}$$

On the other hand, for $\mathbf{e}_0 = -\varepsilon\frac{\mathbf{f}}{\|\mathbf{f}\|}$ which belongs to the feasible set $\|\mathbf{e}\| \leq \varepsilon$, we have $\mathrm{Re}(\mathbf{f}^H\mathbf{e}_0) = -\varepsilon\|\mathbf{f}\|$. This implies that

$$\min_{\|\mathbf{e}\|\leq\varepsilon} \mathrm{Re}(\mathbf{f}^H\mathbf{e}) = -\varepsilon\|\mathbf{f}\|\,. \tag{5.26}$$

Inserting (5.26) into (5.23), we obtain (5.22) and the proof is complete. □

Using (5.21) and (5.22), we can transform (5.20) to

$$\min_{\mathbf{f}}\{\mathbf{f}^H(\hat{\underline{\mathbf{R}}} + \gamma\mathbf{I})\mathbf{f} - \mathbf{f}^H\hat{\underline{\mathbf{d}}} - \hat{\underline{\mathbf{d}}}^H\mathbf{f} + 2\varepsilon A_1\|\mathbf{f}\|\}. \tag{5.27}$$

Differentiating the objective function in (5.27) with respect to $\mathbf{f}^H$, equating it to zero, and using (5.16), we obtain that the solution to the optimization problem (5.27) satisfies the following equation

$$(\hat{\underline{\mathbf{R}}} + \gamma\mathbf{I})\mathbf{f} + \varepsilon A_1\frac{\mathbf{f}}{\|\mathbf{f}\|} = A_1\,\hat{\underline{\mathbf{w}}}_1. \tag{5.28}$$

Note that to solve (5.28) directly, one needs to know A_1. However, this information may be unavailable. To avoid this difficulty, let us rescale the vector $\mathbf{f}$ by the factor of A_1. This operation is motivated by the fact that rescaling $\mathbf{f}$ by an arbitrary constant we do not change the probability of error of any linear multiuser receiver (although, the MSE can be changed dramatically by such a rescaling). Taking into account that our ultimate goal is the probability of error performance and using, for the sake of simplicity, the same notation $\mathbf{f}$ for the rescaled coefficient vector, we can rewrite (5.28) as

$$(\hat{\underline{\mathbf{R}}} + \gamma\mathbf{I})\mathbf{f} + \varepsilon\frac{\mathbf{f}}{\|\mathbf{f}\|} = \hat{\underline{\mathbf{w}}}_1. \tag{5.29}$$

We stress again that the advantage of using (5.29) instead of (5.28) is that the knowledge of A_1 is not required in (5.29).

To solve (5.29), let us rewrite it as

$$(\hat{\underline{\mathbf{R}}} + (\gamma + \varepsilon/\|\mathbf{f}\|)\mathbf{I})\mathbf{f} = \hat{\underline{\mathbf{w}}}_1 \tag{5.30}$$

or, equivalently, as

$$\mathbf{f} = (\hat{\underline{\mathbf{R}}} + (\gamma + \varepsilon/\|\mathbf{f}\|)\mathbf{I})^{-1}\hat{\underline{\mathbf{w}}}_1. \tag{5.31}$$

We observe that the robust multiuser receiver (5.31) uses an adaptive diagonal loading factor $\gamma + \varepsilon/\|\mathbf{f}\|$ which depends on the norm of the receiver coefficient vector $\mathbf{f}$ itself. This factor is optimally matched to known amounts of uncertainty in the user signature and data covariance matrix. It should be mentioned that the robust receiver (5.31) has been extended in [65] to the MIMO space-time coded systems. A related principle as in our technique has also been used in [43] to obtain robust Capon receiver for DOA estimation.

A noteworthy observation following from (5.31) is that, if $\|\mathbf{f}\|$ is available then we can use (5.31) to calculate the coefficient vector of the proposed robust blind receiver. In what follows, we propose a simple method to determine $\|\mathbf{f}\|$.

Taking the norm of the both sides of (5.31), we have

$$\|\mathbf{f}\|^2 = \|(\underline{\hat{\mathbf{R}}} + (\gamma + \varepsilon/\|\mathbf{f}\|)\mathbf{I})^{-1}\underline{\hat{\mathbf{w}}}_1\|^2. \tag{5.32}$$

Let us introduce

$$\tau = \|\mathbf{f}\| > 0. \tag{5.33}$$

We have that solving (5.32) is equivalent to finding a positive value of τ such that

$$\tau^2 = \|(\underline{\hat{\mathbf{R}}} + (\gamma + \varepsilon/\tau)\mathbf{I})^{-1}\underline{\hat{\mathbf{w}}}_1\|^2. \tag{5.34}$$

To simplify (5.34), let us write the eigendecomposition of $\underline{\hat{\mathbf{R}}}$ as

$$\underline{\hat{\mathbf{R}}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \tag{5.35}$$

where $\mathbf{U}$ is the $L_c \times L_c$ unitary matrix whose columns are the eigenvectors of $\underline{\hat{\mathbf{R}}}$ and $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues of $\underline{\hat{\mathbf{R}}}$ given by

$$\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_{L_c}\}. \tag{5.36}$$

Here, $\{\lambda_i\}_{i=1}^{L_c}$ are the real positive eigenvalues of $\underline{\hat{\mathbf{R}}}$. Without loss of generality, we assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{L_c} > 0$. Using (5.35), we can rewrite (5.34) as

$$\|\mathbf{U}\mathbf{\Psi}^{-1}(\tau)\mathbf{U}^H\underline{\hat{\mathbf{w}}}_1\|^2 - \tau^2 = 0 \tag{5.37}$$

where

$$\mathbf{\Psi}(\tau) = \mathbf{\Lambda} + (\gamma + \varepsilon/\tau)\mathbf{I}. \tag{5.38}$$

Introducing the $L_c \times 1$ vector $\breve{\mathbf{w}}$ as

$$\breve{\mathbf{w}} = [\breve{w}_1, \ldots, \breve{w}_{L_c}]^T = \mathbf{U}^H \underline{\hat{\mathbf{w}}}_1 \tag{5.39}$$

and taking into account that $\mathbf{U}$ is a unitary matrix, we can express the left hand side of (5.37) as

$$\begin{aligned} \|\mathbf{U}\boldsymbol{\Psi}^{-1}(\tau)\mathbf{U}^H \underline{\hat{\mathbf{w}}}_1\|^2 - \tau^2 &= \|\boldsymbol{\Psi}^{-1}(\tau)\breve{\mathbf{w}}\|^2 - \tau^2 \\ &= \sum_{i=1}^{L_c} \Big(\frac{|\breve{w}_i|}{\lambda_i + \gamma + \frac{\varepsilon}{\tau}} \Big)^2 - \tau^2 \\ &= \left[\sum_{i=1}^{L_c} \Big(\frac{|\breve{w}_i|}{\varepsilon + \tau(\lambda_i + \gamma)} \Big)^2 - 1 \right] \tau^2. \end{aligned} \tag{5.40}$$

Using (5.40) and noting that $\tau > 0$, we obtain that solving (5.37) is equivalent to finding a positive value for τ such that

$$f(\tau) = \sum_{i=1}^{L_c} \Big(\frac{|\breve{w}_i|}{\varepsilon + \tau(\lambda_i + \gamma)} \Big)^2 - 1 = 0. \tag{5.41}$$

Note that (5.41) may not always have a real and positive solution. The following lemma states the necessary and sufficient conditions under which (5.41) has a unique positive solution.

Lemma 3: Equation (5.41) has a unique real-valued and positive solution if and only if

$$\|\underline{\hat{\mathbf{w}}}_1\| > \varepsilon. \tag{5.42}$$

Proof: We first show that if $\varepsilon < \|\underline{\hat{\mathbf{w}}}_1\|$ then the solution of $f(\tau) = 0$ is a positive value. To show this, we note that

$$\begin{aligned} f(0) &= \frac{\sum_{i=1}^{L_c} |\breve{w}_i|^2}{\varepsilon^2} - 1 \\ &= \frac{\|\breve{\mathbf{w}}\|^2}{\varepsilon^2} - 1 \\ &= \frac{\|\underline{\hat{\mathbf{w}}}_1\|^2}{\varepsilon^2} - 1 \end{aligned} \tag{5.43}$$

where in the last equation we have used the fact that $\|\check{\mathbf{w}}\| = \|\underline{\hat{\mathbf{w}}}_1\|$, which, in turn, follows from (5.39) and the fact that $\mathbf{U}$ is unitary. If $\varepsilon < \|\underline{\hat{\mathbf{w}}}_1\|$, then it follows from (5.43) that $f(0) > 0$. On the other hand, $f(+\infty) = -1$ and, since $f(\tau)$ is continuous for positive values of τ, it has a root in the interval $(0, +\infty)$. This completes the proof of sufficiency.

The necessity of the condition $\varepsilon < \|\underline{\hat{\mathbf{w}}}_1\|$ for $f(\tau) = 0$ to have a positive solution can be shown by contradiction. Assume that the equation $f(\tau) = 0$ has a positive solution while $\varepsilon \geq \|\underline{\hat{\mathbf{w}}}_1\|$. Since τ, γ, and $\{\lambda_i\}_{i=1}^{L_c}$ are all positive, using the definition of $f(\tau)$ in (5.41) we obtain that for any positive τ

$$\begin{aligned} f(\tau) &< \frac{\sum_{i=1}^{L_c} |\check{w}_i|^2}{\varepsilon^2} - 1 \\ &= \frac{\|\check{\mathbf{w}}\|^2}{\varepsilon^2} - 1 \\ &= \frac{\|\underline{\hat{\mathbf{w}}}_1\|^2}{\varepsilon^2} - 1. \end{aligned} \tag{5.44}$$

If $\varepsilon \geq \|\underline{\hat{\mathbf{w}}}_1\|$, it follows from (5.44) that $f(\tau) < 0$ for all positive values of τ. This contradicts to the assumption that $f(\tau)$ is zero for some positive τ and proves the necessity part.

The proof of uniqueness is as follows. Assume that τ_1 and τ_2 are two positive values of τ such that $f(\tau_1) = f(\tau_2)$. Then, using (5.41), we can write

$$\sum_{i=1}^{L_c} \Big(\frac{|\check{w}_i|}{\varepsilon + \tau_1(\lambda_i + \gamma)} \Big)^2 - \sum_{i=1}^{L_c} \Big(\frac{|\check{w}_i|}{\varepsilon + \tau_2(\lambda_i + \gamma)} \Big)^2 = 0 \tag{5.45}$$

which means that

$$(\tau_2 - \tau_1) \sum_{i=1}^{L_c} \frac{|\check{w}_i|^2 (\lambda_i + \gamma)[2\varepsilon + (\lambda_i + \gamma)(\tau_2 + \tau_1)]}{[\varepsilon + \tau_1(\lambda_i + \gamma)]^2 [\varepsilon + \tau_2(\lambda_i + \gamma)]^2} = 0 \tag{5.46}$$

where

$$\sum_{i=1}^{L_c} \frac{|\check{w}_i|^2 (\lambda_i + \gamma)[2\varepsilon + (\lambda_i + \gamma)(\tau_2 + \tau_1)]}{[\varepsilon + \tau_1(\lambda_i + \gamma)]^2 [\varepsilon + \tau_2(\lambda_i + \gamma)]^2} > 0. \tag{5.47}$$

This means that $\tau_1 = \tau_2$ and, therefore, the solution to $f(\tau) = 0$ is unique. This completes the proof. □

The condition (5.42) has a simple and intuitively appealing interpretation. Note that the parameter ε characterizes the maximal norm of the error between the estimated and the actual user signatures. Therefore, (5.42) simply states that our approach is applicable only if the maximum norm of such an error is small enough, so that it does not exceed the norm of the estimated user signature itself. In what follows, we assume that (5.42) is always satisfied.

Using (5.41), we can upper-bound the function $f(\tau)$ as

$$\begin{aligned} f(\tau) &< \frac{\sum_{i=1}^{L_c} |\breve{w}_i|^2}{(\varepsilon + \tau(\lambda_{L_c} + \gamma))^2} - 1 \\ &= \frac{\|\breve{\mathbf{w}}\|^2}{(\varepsilon + \tau(\lambda_{L_c} + \gamma))^2} - 1 \\ &= \frac{\|\hat{\underline{\mathbf{w}}}_1\|^2}{(\varepsilon + \tau(\lambda_{L_c} + \gamma))^2} - 1 = f_{\text{up}}(\tau). \end{aligned} \tag{5.48}$$

Noting that $f(\tau)$ and $f_{\text{up}}(\tau)$ are both decreasing functions for positive values of τ and that, according to Lemma 3, the root τ of $f(\tau)$ is positive, we obtain from (5.48) that this root is always smaller than the root

$$\tau_{\text{up}} = \frac{\|\hat{\underline{\mathbf{w}}}_1\| - \varepsilon}{\lambda_{L_c} + \gamma} \tag{5.49}$$

of $f_{\text{up}}(\tau)$. Hence, the value of τ belongs to the interval $(0, \tau_{\text{up}})$. Note that with this condition, the problem of finding τ becomes standard. For example, the algorithm of [103] can be used for this purpose. This algorithm consists of a binary search followed by Newton-Raphson iterations. The binary search technique is used to obtain a proper initial point for the subsequent Newton-Raphson procedure. As shown in [103], this algorithm converges to a ζ-neighborhood of τ in $\mathcal{O}\left(\log\log(\tau_{\text{up}}/\zeta)\right)$ iterations. The algorithm to obtain $\|\mathbf{f}\|$ can be summarized as follows:

1. Use binary search to find $\tau_0 \in (0, \tau_{\text{up}})$ such that $f(\tau_0) > 0$ and $f(\frac{13}{12}\tau_0) < 0$ (see [103] for details).

2. Set $l = 1$ and select a small positive value of ξ which will be used in the algorithm stopping criterion.

3. Obtain τ_l as
$$\tau_l = \tau_{l-1} - \frac{f(\tau_{l-1})}{f'(\tau_{l-1})} \tag{5.50}$$
where $f'(\tau_{l-1})$ is the derivative of $f(\tau)$ at $\tau = \tau_{l-1}$.

4. If $|f(\tau_l)| < \xi$, go to the next step. Otherwise, repeat steps 2 and 3.

5. Determine $\|\mathbf{f}\|$ as $\tau = \tau_l$.

5.3.2 Summary of the proposed multiuser receiver

As the procedure developed in the previous subsection enables us to find the value of $\|\mathbf{f}\|$, it can be directly used to compute the coefficient vector of our multiuser receiver. Using this fact, the proposed multiuser detection algorithm can be summarized as follows:

1. Compute the sample covariance matrix $\hat{\underline{\mathbf{R}}}$ from the received data.

2. Find the eigendecomposition (5.35) of $\hat{\underline{\mathbf{R}}}$.

3. Compute $\breve{\mathbf{w}}$ using (5.39).

4. Find the value of $\tau = \|\mathbf{f}\|$ using the Newton-Raphson procedure summarized in the previous subsection.

5. Compute the coefficient vector $\mathbf{f}_{\text{rob}}$ of the proposed robust blind multiuser receiver as
$$\mathbf{f}_{\text{rob}} = (\hat{\underline{\mathbf{R}}} + (\gamma + \varepsilon/\tau)\mathbf{I})^{-1}\underline{\hat{\mathbf{w}}}_1. \tag{5.51}$$

The dominant computational cost of the summarized algorithm is determined by that of the eigendecomposition in Step 2 and matrix inversion in Step 5.

5.3.3 Relationship to the SOCP multiuser receiver

In this section, we show the relationship between our robust multiuser receiver (5.51) and the SOCP-based robust multiuser receiver originally proposed in [16] for the real-valued case and extended in [86] to the complex-valued case[3].

[3]Note that in [86], the adaptive beamforming problem rather than the multiuser detection problem is considered.

First of all, let us briefly revisit the formulation of [86]. It has been shown in this paper that a robust MOE-based receiver based on the worst-case performance optimization can be designed by solving the following optimization problem

$$\min_{\mathbf{f}} \mathbf{f}^H \hat{\underline{\mathbf{R}}} \mathbf{f} \quad \text{subject to} \quad |\mathbf{f}^H \hat{\underline{\mathbf{w}}}_1| - \varepsilon \|\mathbf{f}\| \geq 1. \tag{5.52}$$

The authors of [86] converted (5.52) to the SOCP problem. Below, we will show that the algorithm presented in the previous section can be easily used to solve (5.52). Let us first prove that if the objective function in (5.52) is minimized, then the inequality constraint in (5.52) is satisfied as equality. This can be proved by contradiction. Assume that the objective function $\mathbf{f}^H \hat{\underline{\mathbf{R}}} \mathbf{f}$ in (5.52) is minimized at $\check{\mathbf{f}}$ for which the constraint in (5.52) is satisfied as a strict inequality $|\check{\mathbf{f}}^H \hat{\underline{\mathbf{w}}}_1| - \varepsilon \|\check{\mathbf{f}}\| = \alpha > 1$. Then, introducing a new coefficient vector $\overline{\mathbf{f}} = \check{\mathbf{f}}/\alpha$, we obtain that with such $\overline{\mathbf{f}}$, the constraint in (5.52) is still satisfied, while the objective function is further reduced. This is an obvious contradiction because it has been assumed that the cost function has already achieved its minimum. Hence, $\alpha = 1$ which means that the problem (5.52) can be rewritten as

$$\min_{\mathbf{f}} \mathbf{f}^H \hat{\underline{\mathbf{R}}} \mathbf{f} \quad \text{subject to} \quad |\mathbf{f}^H \hat{\underline{\mathbf{w}}}_1| - \varepsilon \|\mathbf{f}\| = 1. \tag{5.53}$$

Next, we observe that the cost function in (5.53) is unchanged when $\mathbf{f}$ undergoes an arbitrary phase rotation [86]. Therefore, we can always rotate, without affecting the objective function value, the phase of $\mathbf{f}$ so that $\mathbf{f}^H \hat{\underline{\mathbf{w}}}_1$ is real. Thus, without loss of generality, we can rewrite (5.53) as

$$\min_{\mathbf{f}} \mathbf{f}^H \hat{\underline{\mathbf{R}}} \mathbf{f} \quad \text{subject to} \quad \mathbf{f}^H \hat{\underline{\mathbf{w}}}_1 - \varepsilon \|\mathbf{f}\| = 1. \tag{5.54}$$

The solution to (5.54) can be obtained using the Lagrange multiplier method. The Lagrangian function can be written as

$$L(\mathbf{f}, \mu) = \mathbf{f}^H \hat{\underline{\mathbf{R}}} \mathbf{f} - \mu(\mathbf{f}^H \hat{\underline{\mathbf{w}}}_1 - \varepsilon \|\mathbf{f}\| - 1) \tag{5.55}$$

where μ is the Lagrange multiplier. Differentiating $L(\mathbf{f}, \mu)$ with respect to $\mathbf{f}^H$ and equating the result to zero, we obtain the following equation:

$$\hat{\underline{\mathbf{R}}} \mathbf{f} + \mu \varepsilon \frac{\mathbf{f}}{\|\mathbf{f}\|} = \mu \hat{\underline{\mathbf{w}}}_1. \tag{5.56}$$

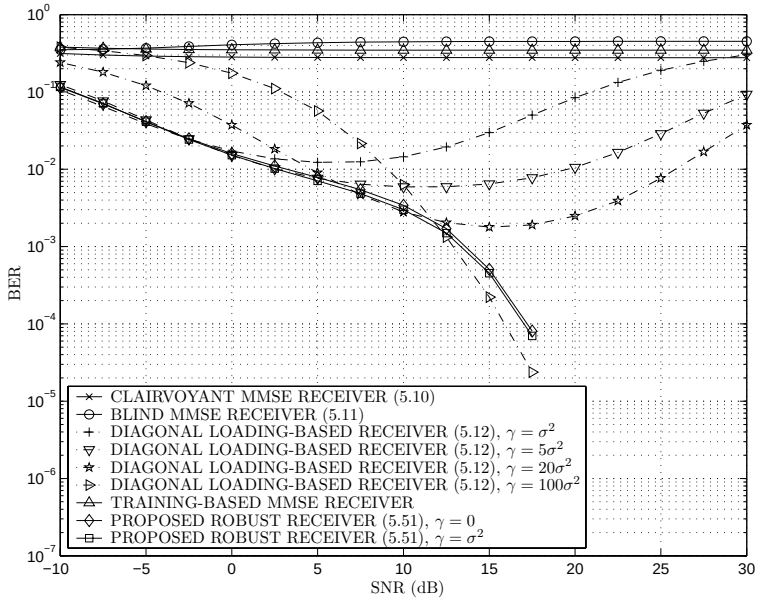

Figure 5.1: BERs versus SNR. First example.

To solve (5.56), one needs to know the Lagrange multiplier μ. However, using the fact that multiplying $\mathbf{f}$ in (5.56) by any arbitrary constant does not change the probability of error at the output of the symbol detector, we can transform this equation as

$$\hat{\underline{\mathbf{R}}}\mathbf{f} + \varepsilon \frac{\mathbf{f}}{\|\mathbf{f}\|} = \hat{\underline{\mathbf{w}}}_1. \tag{5.57}$$

As before, for the sake of simplicity the same notation $\mathbf{f}$ is used in (5.57) for the rescaled coefficient vector as for the original one in (5.56). Comparing (5.57) with (5.29), we obtain that if $\gamma = 0$ in (5.29) then it transforms to (5.57). Therefore, the MOE-based receiver of [16] and its extension of [86] are special cases of our MMSE-based multiuser receiver. This also implies that the receivers of [16] and [86] can be implemented using the algorithm developed above rather than the SOCP algorithm.

5.4 Simulations

In our simulation examples, we consider a 7-user synchronous CDMA system. The BPSK modulation scheme is used and binary Gold codes of the length $L_c = 31$ are employed as

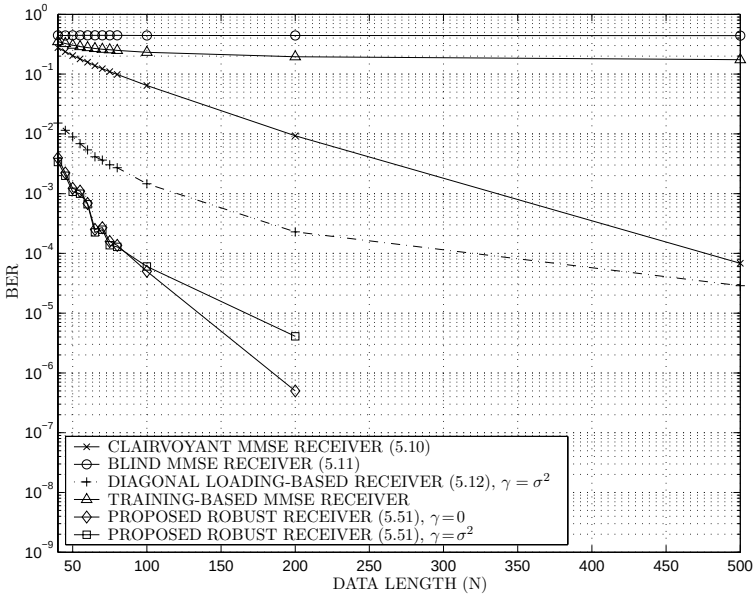

Figure 5.2: BERs versus N. First example.

user spreading codes. Throughout all our examples, all the interfering users are assumed to have the INR equal to 20 dB. The performances of the following multiuser receivers are compared in terms of the BER at the output of the symbol detector:

- the clairvoyant multiuser receiver (5.10) which corresponds to the ideal case when the user signature $\underline{\mathbf{w}}_1$ is known exactly (this algorithm does not correspond to any practical situation but is considered for the reason of comparison only).
- the blind multiuser receiver (5.11).
- the diagonal loading-based blind multiuser receiver (5.12) with different ad hoc values of the diagonal loading factor γ.
- the training-based MMSE multiuser receiver.
- the proposed blind multiuser receiver (5.51).

In the training-based multiuser receiver, 30 samples are used to estimate $\underline{\mathbf{d}}$. A total of 1000 runs is used to obtain each point of the BER curves.

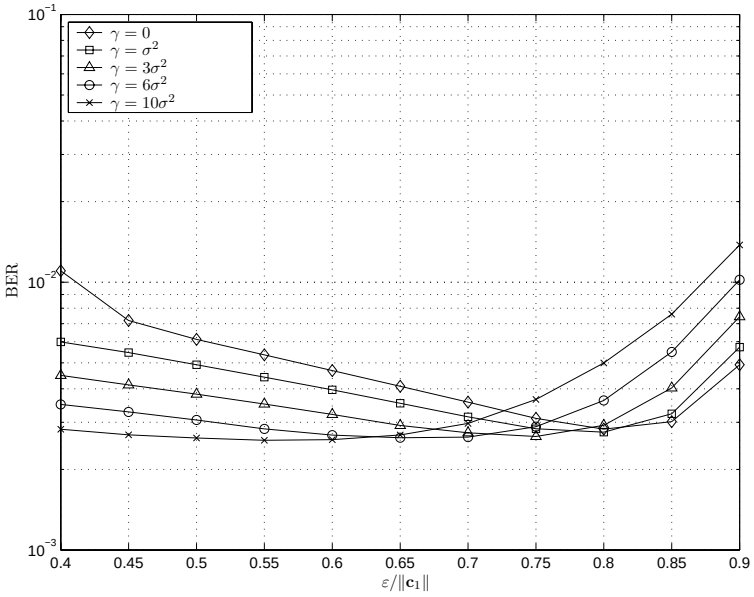

Figure 5.3: BERs of the proposed receiver versus $\varepsilon/\|\mathbf{c}_1\|$. First example.

In the first channel model used in Figs. 5.1-5.4, each of the user spreading codes is distorted by an additive random Gaussian vector drawn uniformly from the interval $[-\delta, \delta]$. For each user, such a random vector is added to the spreading code vector to simulate the effect of ICI [16]. Ignoring the effect of unknown channel, the blind receivers (5.11), (5.12) and (5.51) use the spreading code vector $\mathbf{c}_1$ as the presumed user signature $\underline{\hat{\mathbf{w}}}_1$ [16], [31], [75], [89], in Figs. 5.1-5.4. In these figures, the upper bound for the norm of the error vector $\mathbf{e}$ is equal to $\delta\sqrt{L_c}$, and hence, $\varepsilon = \delta\sqrt{L_c}$ has been chosen.

Fig. 5.1 shows the BERs of the multiuser receivers tested versus the SNR of the user-of-interest. In this figure, $N = 40$ data vectors are used to obtain the sample covariance matrix $\underline{\hat{\mathbf{R}}}$, and $\delta = 0.7$ is chosen. Note that this choice of δ implies that the amount of ICI per chip is up to 70%. As it can be seen from Fig. 5.1, our robust multiuser receiver provides the best performance trade-off over all SNR values. It can be seen that the clairvoyant multiuser receiver shows a poor performance which is due to the short data length effect. Furthermore, the BER of the diagonal loading-based multiuser receiver does not decrease monotonically when the SNR increases. If the diagonal loading factor comparable to the

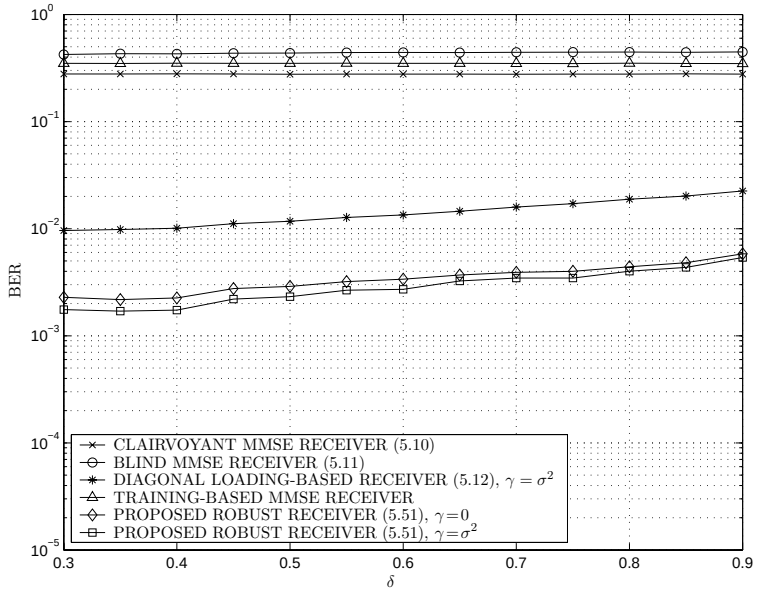

Figure 5.4: BERs versus δ. First example.

noise power σ^2 is chosen, then the performance of the conventional diagonal loading-based multiuser receiver (5.12) becomes poor at high SNRs. Vice versa, if $\gamma \gg \sigma^2$ is chosen, then the performance of the latter receiver becomes poor at low SNRs. Note that our robust multiuser receiver (5.51) uses an adaptive diagonal loading factor whose value varies with the SNR and is optimally matched to the uncertainties in the presumed signature of the user-of-interest and data covariance matrix. This explains why the performance of (5.51) is good over a wide range of SNR. Interestingly, our robust receiver has excellent performance even when $\gamma = 0$. The explanation of this fact follows from equation (5.31) which shows that both γ and ε contribute to the adaptive diagonal loading of the sample covariance matrix, i.e., if $\gamma = 0$ then robustness is provided by means of ε.

Fig. 5.2 shows the BERs of the multiuser receivers tested versus the data length N. In this figure SNR = 10 dB, $\delta = 0.7$ and $\varepsilon = \delta\sqrt{L_c}$ are chosen. As it can be seen from this figure, the proposed multiuser receiver has substantially faster convergence rate as compared to the other multiuser detection techniques.

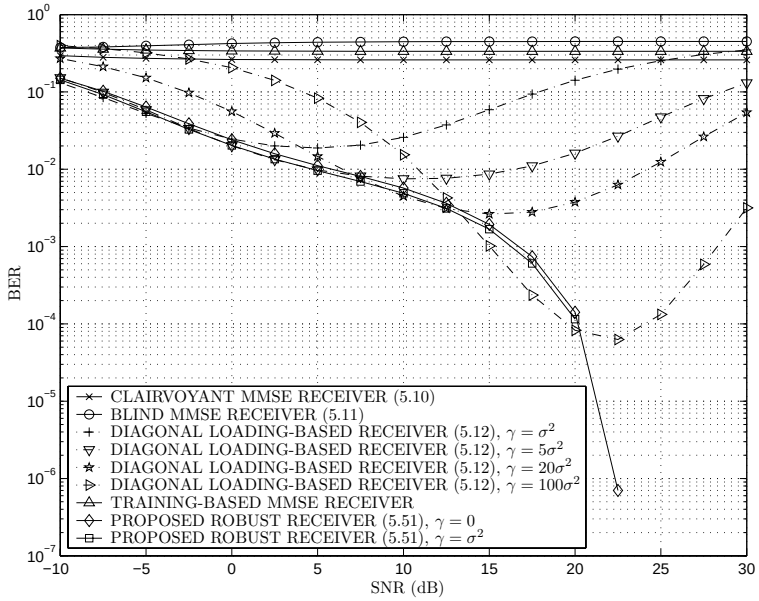

Figure 5.5: BERs versus SNR. Second example.

To study the effect of selection of the parameters ε and γ, the BER of our robust multiuser receiver is shown in Fig. 5.3 versus $\varepsilon/\|\mathbf{c}_1\| = \varepsilon/\sqrt{L_c}$ for different values of γ. In this figure, SNR = 10 dB, $\delta = 0.7$ and $N = 40$ are chosen. As it can be seen from Fig. 5.3, when γ is comparable to the noise power σ^2, the performance of the proposed method is less sensitive to the selection of ε compared to the case when $\gamma = 0$ or when γ is much larger than σ^2.

To study the effect of an improper choice of the parameter δ, the BERs of the multiuser receivers tested are shown versus δ in Fig. 5.4. In this figure, we take SNR = 10 dB, $N = 40$, and $\varepsilon/\sqrt{L_c} = 0.7$. Such a choice of ε implies that we assume δ to be equal to 0.7 while the actual value of δ varies between 0.3 and 0.9. As can be seen from Fig. 5.4, overestimating δ can even slightly improve the performance of the proposed multiuser receiver while underestimating δ does not affect the performance significantly. This slight performance improvement can be explained by the fact that the proposed worst-case design is quite conservative and even less conservative designs may be sufficient for providing the robustness required.

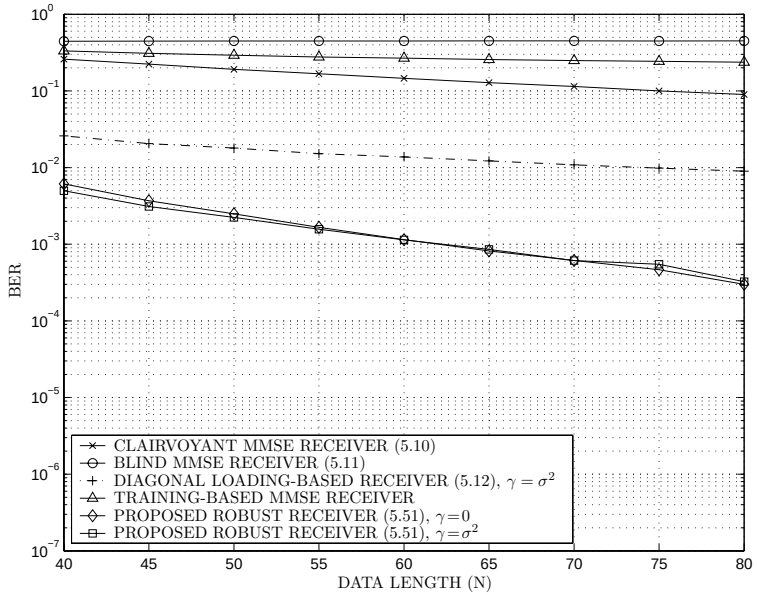

Figure 5.6: BERs versus N. Second example.

In the second example, we model the channel with a vector of length $L = 4$. Similar to the previous figures, we ignore the effect of the unknown channel and use $\mathbf{c}_1$ instead of the user signature estimate $\underline{\hat{\mathbf{w}}}_1$.

For each simulation run, the channel is modeled as $\mathbf{h} = [\delta_0, \delta_1 e^{j\phi_1}, 0, \delta_3 e^{j\phi_3}]^T$ where δ_0, δ_1, and δ_3 are, respectively, the amplitudes of the first, the second and the fourth entry of $\mathbf{h}$ whereas ϕ_1 and ϕ_3 are the phases of the second and the fourth entry, respectively. In each simulation run, δ_0 is randomly chosen in the interval $[0.9, 1.1]$, δ_1 and δ_3 are uniformly distributed over the interval $[0.45, 0.55]$, while ϕ_1 and ϕ_3 are randomly drawn from a uniform distribution over the interval $[0, 2\pi]$.

Fig. 5.5 displays the BERs of the multiuser receivers tested versus the SNR of the user-of-interest. In this figure, $N = 40$ data vectors are used to obtain the sample covariance matrix $\hat{\mathbf{R}}$. In our robust multiuser receiver, $\varepsilon = 0.7\|\mathbf{c}_1\| = 0.7\sqrt{L_c}$ is chosen. As it can be seen from Fig. 5.5, the proposed robust multiuser receiver provides the best performance trade-off among the multiuser receivers tested over the wide SNR range. There is an obvious similarity in the receiver performance curves with Fig. 5.1.

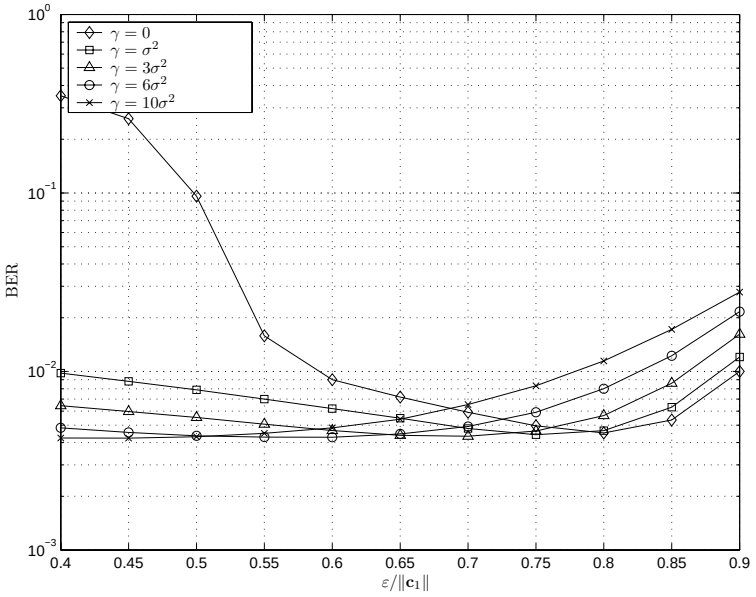

Figure 5.7: The BER of the proposed receiver versus $\varepsilon/\sqrt{L_c}$. Second example.

In Fig. 5.6, the BERs of the multiuser receivers tested are displayed versus the data length N. In this figure, SNR = 10 dB and $\varepsilon = 0.7\sqrt{L_c}$ is chosen. As one can see from Fig. 5.6, the proposed robust receiver converges much faster than the conventional diagonal loading-based receiver.

Fig. 5.7 shows the BER of the proposed multiuser receiver versus $\varepsilon/\sqrt{L_c}$ for different values of the diagonal loading factor γ. Similar to Fig. 5.3, for γ comparable to σ^2 the proposed robust multiuser receiver is less sensitive to the selection of ε as compared to the case when very small ($\gamma \simeq 0$) or very large ($\gamma \gg \sigma^2$) values of γ are used.

In our last example, we test the blind methods (5.11), (5.12) and (5.51) where $\underline{\hat{\mathbf{w}}}_1$ is the blind subspace-based estimate of the signature of the user-of-interest which is obtained from 30 data vectors using the LX algorithm [47]. The same scenario and channel distortion model are chosen as in our second example. Fig. 5.8 displays the BERs of the multiuser receivers tested versus the SNR of the user-of-interest. It can be seen that even if the blind estimate $\underline{\hat{\mathbf{w}}}_1$ is the proposed robust multiuser receivers perform much better than the other techniques tested.

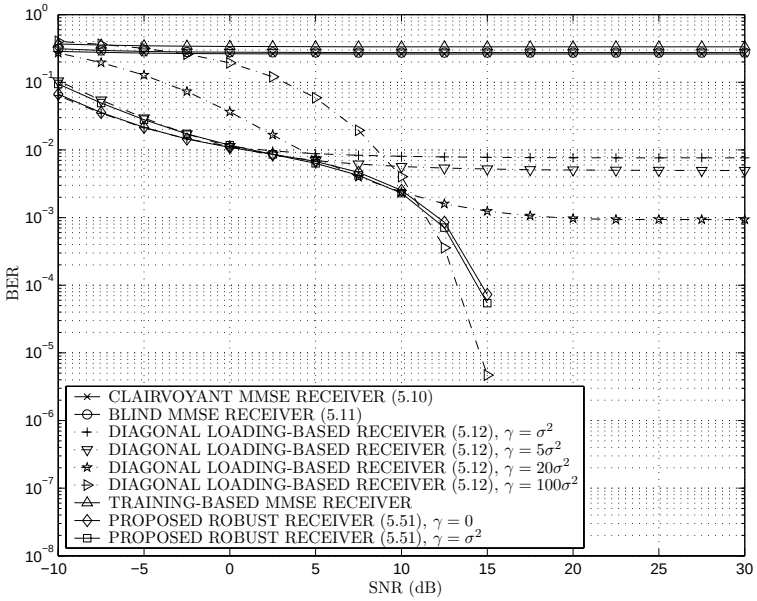

Figure 5.8: BERs versus SNR. Third example.

5.5 Chapter summary

In this chapter, a new blind multiuser receiver has been proposed which is robust against the effects of an erroneous presumed signature of the user-of-interest and short data length. Our approach is based on the explicit modeling of possible mismatches in the MSE cost function and worst-case performance optimization. We have shown that the proposed approach results into a new multiuser receiver structure which uses the sample data covariance matrix with an adaptive diagonal loading. A computationally efficient algorithm has been developed to obtain the receiver coefficient vector. Simulation results have shown substantial performance improvements achieved by our approach relative to the existing training-based and blind multiuser detection techniques.

Chapter 6

Asymptotic performance analysis of MOE receivers in large DS-CDMA systems

6.1 Introduction

As mentioned in the previous chapters, linear blind multiuser receivers represent an attractive choice of multiuser detection techniques since they effectively mitigate MAI while maintaining high spectral efficiency and low computational complexity [31], [72], [84], [91], [117]. The MOE approach has recently emerged as a promising trend among such linear blind techniques [31], [42], [75], [78], [93], [101]. MOE multiuser receivers aim at suppressing MAI while preserving the energy of the user-of-interest at the receiver output by means of a set of constraints guaranteeing that the energy of the user-of-interest is held constant. These constraints can be treated as design parameters that should be properly chosen to improve the receiver performance. Borrowing the idea of the Capon estimation technique [12] from array processing to maximize the energy of the user-of-interest after interference suppression, one can select these constraints so that the resulting MOE is maximized [32],

[41], [42], [75], [78], [100]. The constraint vector and the MOE receiver obtained from such max-min procedure are known as the *Capon channel estimate* and the *Capon receiver*, respectively [32], [42], [75]. It has been shown in [78] that under a certain identifiability condition, the Capon channel estimate converges to a scaled version of the user-of-interest channel vector as the noise power converges to zero.

In spite of extensive research on designing advanced MOE receivers, the available analytical results on the performance of these receivers are insufficient. This is especially true regarding the receiver performance analysis in *large random communication systems* wherein the user spreading codes are modelled as random vectors and the number of users goes to infinity, while the system load (that is, the ratio of the number of users to the spreading factor) remains constant. Such approach to the receiver performance analysis has been originally proposed in [79] and [85] to analyze the MMSE and decorrelating receivers for synchronous DS-CDMA systems with flat fading channels. In these papers, it has been shown that in such large random communication systems, the receiver SINR becomes independent from the realization of the spreading code. This key observation made it possible to derive important results on the receiver performance [79], [85] that cannot be obtained using the conventional approaches. These results are of great practical interest as they also approximately hold in finite-size systems with non-random spreading codes. Motivated by this fact, there has been a surge of research on using the approach of [79] and [85] to receiver performance analysis. The results of [79] has been generalized to the asynchronous and multi-antenna DS-CDMA systems in [35] and [29], respectively. The downlink performance of the MMSE receiver in a multipath channel has been studied in [14] and the performance of the MMSE receiver conjoint with a channel estimation technique has been analyzed in [19]. Other relevant applications can be found in [17] and [59]. This approach has also been used to analyze and design of the reduced-rank MMSE receivers [28], [33], [44] and large MIMO communication systems [18], [52], [55], [80].

In spite of the increasing popularity of technique introduced in [79], [85], to the best of our knowledge, there have been no attempts to analyze the blind MOE and Capon receivers using such an approach. In this chapter (see, also [107], [113]), we use this approach to

analyze the performances of the MOE and Capon receivers assuming that the user spreading codes are random i.i.d. vectors, and both the number of users and the spreading factor go to infinity with the same rate. The results of this chapter correspond to an arbitrary value of SNR, while the asymptotic properties of the MOE receiver at high SNRs will be studied in Chapter 7.

We first adopt a conventional simplifying assumption that the delayed versions of the i.i.d. spreading codes are statistically independent [19], [59]. Using such an assumption, we obtain the asymptotic SINR of the MOE receiver. It is shown that if the empirical distribution of the received user powers converges to a nonrandom distribution, then the asymptotic SINR also converges to a deterministic value that is independent from the particular realization of the spreading codes. It is also shown that the Capon channel estimate converges to a scaled version of the actual channel vector of the user-of-interest, and, moreover, the difference between the SINRs of the Capon receiver and the optimal MMSE receiver[1] converges to zero.

Numerical examples show that the asymptotic results using the aforementioned assumption on statistical independence of the delayed versions of the spreading codes can accurately predict the behavior of the MOE receiver in many practical scenarios. However, in some cases the discrepancies between our asymptotic and numerical results are quite substantial, which implies that the latter assumption is not always relevant. This motivates us to derive more general asymptotic expressions for the SINR of the MOE receiver and Capon channel estimate that do not use this assumption. We obtain such expressions for i.i.d. circular Gaussian spreading codes [14], [19], [28] and show that if the users have either single-path or multipath channels with zero mean and independent channel coefficients, then these expressions transfer to our earlier expressions obtained under the assumption of statistical independence of delayed versions of the spreading codes. In particular, if the user channel coefficients are zero-mean and independent, then the Capon channel estimate and the SINR of the Capon receiver converge to the channel vector of the user-of-interest and the SINR

[1]The optimal MMSE receiver uses the full channel knowledge of the user-of-interest to achieve the maximum possible SINR [79], [50].

of the optimal MMSE receiver, respectively. As the assumption of zero-mean independent channel vectors is relevant for the uplink mode but is violated in the downlink mode, it can be concluded that the Capon receiver and the Capon channel estimation technique are more useful for the uplink rather than for the downlink mode.

This chapter is organized as follows. Section 6.2 gives a brief overview on the MOE receiver. Section 6.3 presents our main results on the asymptotic performance of the MOE receiver and the Capon channel estimator. The effects of the channel vector distribution on the performance of the MOE receiver and the Capon channel estimator are discussed in Section 6.4. Section 6.5 presents simulation results. A summary of the chapter is given in Section 6.6.

6.2 Background

In this chapter, we keep the assumption that the first user is the user-of-interest and, similar to Chapter 5, we consider the linear receivers whose design problem is to determine a vector $\mathbf{f}$ such that the detection of $b_1(n)$ can be made from the receiver output[2]

$$y(n) = \mathbf{f}^H \underline{\mathbf{x}}(n). \tag{6.1}$$

The MOE approach to the receiver design is to minimize the output energy

$$\mathrm{E}\{|y(n)|^2\} = \mathbf{f}^H \underline{\mathbf{R}} \mathbf{f} \tag{6.2}$$

subject to one or more constraints guaranteeing that the energy of the user-of-interest at the receiver output is held constant.

In this and the next chapter, we consider the white noise scenario. Moreover, we assume for the sake of simplicity and without any loss of generality that the signal amplitudes $A_k,\ k = 1, \ldots, K$ are absorbed in the corresponding channel vectors $\mathbf{h}_k,\ k = 1, \ldots, K$.

[2]the receiver output has been earlier introduced in (5.1), but is repeated in (6.1) for the sake of clarity.

Using this assumption along with the fact that $\underline{\mathbf{\Sigma}}_v = \sigma^2\mathbf{I}$, it follows from (1.17) that

$$\begin{aligned}\underline{\mathbf{R}} &= \sum_{k=1}^{K} \boldsymbol{\mathcal{H}}_k \mathbf{c}_k \mathbf{c}_k^H \boldsymbol{\mathcal{H}}_k^H + \sum_{k=1}^{K} \tilde{\boldsymbol{\mathcal{H}}}_k \mathbf{c}_k \mathbf{c}_k^H \tilde{\boldsymbol{\mathcal{H}}}_k^H + \sigma^2\mathbf{I} \\ &= \sum_{k=1}^{K} \underline{\boldsymbol{\mathcal{C}}}_k \mathbf{h}_k \mathbf{h}_k^H \underline{\boldsymbol{\mathcal{C}}}_k^H + \sum_{k=1}^{K} \underline{\tilde{\boldsymbol{\mathcal{C}}}}_k \mathbf{h}_k \mathbf{h}_k^H \underline{\tilde{\boldsymbol{\mathcal{C}}}}_k^H + \sigma^2\mathbf{I}. \end{aligned} \tag{6.3}$$

The MOE receiver that maximizes the SINR

$$\mathrm{SINR}(\mathbf{f}) = \frac{|\mathbf{f}^H \underline{\mathbf{w}}_1|^2}{\mathbf{f}^H \underline{\mathbf{R}} \mathbf{f} - |\mathbf{f}^H \underline{\mathbf{w}}_1|^2} \tag{6.4}$$

is the solution to

$$\min_{\mathbf{f}} \mathbf{f}^H \underline{\mathbf{R}} \mathbf{f} \qquad \text{s.t.} \qquad \mathbf{f}^H \underline{\mathbf{w}}_1 = 1. \tag{6.5}$$

The solution to the latter problem is explicitly given by [32], [78]

$$\mathbf{f}_{\mathrm{opt}} = \frac{1}{\underline{\mathbf{w}}_1^H \underline{\mathbf{R}}^{-1} \underline{\mathbf{w}}_1} \underline{\mathbf{R}}^{-1} \underline{\mathbf{w}}_1. \tag{6.6}$$

Note that $\mathbf{f}_{\mathrm{opt}}$ is a scaled version of the linear MMSE receiver [84]

$$\mathbf{f}_{\mathrm{MMSE}} = \underline{\mathbf{R}}^{-1} \underline{\mathbf{w}}_1. \tag{6.7}$$

Hence, $\mathrm{SINR}(\mathbf{f}_{\mathrm{opt}})$ is equal to [14], [19]

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) = \mathbf{h}_1^H \underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}_1 \tag{6.8}$$

where

$$\mathbf{A} = \underline{\mathbf{R}} - \underline{\mathbf{w}}_1 \underline{\mathbf{w}}_1^H = \underline{\mathbf{R}} - \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}_1 \mathbf{h}_1^H \underline{\boldsymbol{\mathcal{C}}}_1^H. \tag{6.9}$$

The constrained optimization problem (6.5) requires the explicit knowledge of $\underline{\mathbf{w}}_1$. Therefore, in the blind mode (where $\mathbf{h}_1$ and $\underline{\mathbf{w}}_1$ are not known to the receiver), one should resort to some other approaches to obtain the receiver vector. As $\underline{\mathbf{w}}_1 = \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}_1$, an intuitive approach to get around this problem is to modify (6.5) as [32], [78]

$$\min_{\mathbf{f}} \mathbf{f}^H \underline{\mathbf{R}} \mathbf{f} \qquad \text{s.t.} \qquad \mathbf{f}^H \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h} = 1 \tag{6.10}$$

where $\mathbf{h}$ is now an unknown parameter vector. The solution to (6.10) is given by [32]

$$\mathbf{f}_{\mathrm{s}}(\mathbf{h}) = \frac{1}{\mathbf{h}^H \underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}} \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h} \tag{6.11}$$

where the subscript "s" indicates that (6.10) is a single-constraint problem. The corresponding MOE becomes

$$\mathrm{MOE}(\mathbf{f}_{\mathrm{s}}(\mathbf{h})) = \mathbf{f}_{\mathrm{s}}(\mathbf{h})^H \underline{\mathbf{R}} \mathbf{f}_{\mathrm{s}}(\mathbf{h}) = (\mathbf{h}^H \underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h})^{-1}. \tag{6.12}$$

Using the idea of the Capon estimation technique [12], $\mathbf{h}$ can be selected as [32], [78]

$$\mathbf{h}_{\mathrm{C}} = \arg \min_{\mathbf{h}} \mathbf{h}^H \underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}. \qquad \text{s.t.} \qquad \|\mathbf{h}_{\mathrm{C}}\| = 1 \tag{6.13}$$

where the constraint in (6.13) is to avoid the trivial solution $\mathbf{h}_{\mathrm{C}} = \mathbf{0}$. The solution to (6.13) is given by [32]

$$\mathbf{h}_{\mathrm{C}} = \mho \left(\underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \right) = \Omega \left((\underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} \right). \tag{6.14}$$

From (6.3) and the matrix inversion lemma, we have

$$(\underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} = (\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} + \mathbf{h}_1 \mathbf{h}_1^H. \tag{6.15}$$

Hence, (6.14) can be rewritten as

$$\mathbf{h}_{\mathrm{C}} = \Omega \left((\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} + \mathbf{h}_1 \mathbf{h}_1^H \right). \tag{6.16}$$

Substituting (6.16) into (6.11), the Capon receiver vector becomes

$$\mathbf{f}_{\mathrm{C}} = \mathbf{f}_{\mathrm{s}}(\mathbf{h}_{\mathrm{C}}) = \nu \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}_{\mathrm{C}} \tag{6.17}$$

where $\nu = \lambda_{\max} \left((\underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\mathbf{R}}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} \right)$.

Another popular approach to blind MOE receiver design is to obtain the receiver vector from the following optimization problem [32], [75], [78]:

$$\min_{\mathbf{f}} \mathbf{f}^H \underline{\mathbf{R}} \mathbf{f} \qquad \text{s.t.} \qquad \underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{f} = \bar{\mathbf{h}} \tag{6.18}$$

where the constraints in (6.18) preserve the user-of-interest power at the receiver output, and the constraint vector $\bar{\mathbf{h}}$ is a design parameter.

The solution to (6.18) is given by [78]

$$\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}) = \underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1(\underline{\boldsymbol{\mathcal{C}}}_1^H\underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}}. \tag{6.19}$$

where the subscript "m" stresses that the problem in (6.18) is multi-constraint.

From (6.19) it follows that

$$\mathrm{MOE}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = \bar{\mathbf{h}}^H(\underline{\boldsymbol{\mathcal{C}}}_1^H\underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}} \tag{6.20}$$

and

$$\begin{aligned}\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) &= \frac{1}{\dfrac{\bar{\mathbf{h}}^H(\underline{\boldsymbol{\mathcal{C}}}_1^H\underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}}}{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2} - 1} \\ &= \frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\bar{\mathbf{h}}^H(\underline{\boldsymbol{\mathcal{C}}}_1^H\mathbf{A}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}}}\end{aligned} \tag{6.21}$$

where the second equality in (6.21) follows from (6.15). Using an approach similar to (6.13), the MOE in (6.20) can be maximized by selecting $\bar{\mathbf{h}} = \mathbf{h}_{\mathrm{C}}$ which is the solution to the following problem [32], [78]

$$\max_{\bar{\mathbf{h}}} \bar{\mathbf{h}}^H(\underline{\boldsymbol{\mathcal{C}}}_1^H\underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}} \qquad \text{s.t.} \qquad \|\bar{\mathbf{h}}\| = 1. \tag{6.22}$$

Then, the resulting receiver vector becomes

$$\begin{aligned}\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}) &= \underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1(\underline{\boldsymbol{\mathcal{C}}}_1^H\underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\mathbf{h}_{\mathrm{C}} \\ &= \nu\underline{\mathbf{R}}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1\mathbf{h}_{\mathrm{C}} \\ &= \mathbf{f}_{\mathrm{C}}\end{aligned} \tag{6.23}$$

where the second equality in (6.23) follows from (6.14). The above discussion implies that the Capon receiver vector (6.17) is a special form of the multi-constrained MOE receiver vector (6.19) for $\bar{\mathbf{h}} = \mathbf{h}_{\mathrm{C}}$. Hence, throughout the rest of this chapter we analyze the performance of the general MOE receiver (6.19) and its special case (6.23).

6.3 Asymptotic analysis

In this section, we analyze the performances of the Capon channel estimate (6.14) and the MOE receivers (6.19) and (6.23) when the user spreading codes are i.i.d. random vectors and both K and L_c go to infinity with the same rate. The following common assumption will be used throughout the chapter [17], [19], [28], [33], [59]:

A4: There exists a positive constant S such that $\max_k \|\mathbf{h}_k\|^2 \leq S$.

Denoting

$$h_k(z) = \sum_{q=0}^{L-1} h_{k,q} z^{-q} \tag{6.24}$$

we have that **A4** implies that for any k and n

$$|h_k\big(e^{j\frac{2\pi}{L_c}(n-1)}\big)|^2 \leq LS = B. \tag{6.25}$$

Note that as L_c and K go to infinity with the same rate, L is assumed to remain constant [14], [28]. It can be observed from (6.8), (6.16), and (6.21) that the asymptotic performance of the MOE receivers (6.19) and (6.23) as well as that of the MMSE receiver (see also [14], [19], [79]) are essentially dictated by the asymptotic behavior of the matrix $\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1$. To proceed with our analysis, we need to establish the following result.

Proposition 1: Assume that $\mathbf{c}_k = \frac{1}{\sqrt{L_c}} \check{\mathbf{c}}_k$ where the entries of $\check{\mathbf{c}}_k$ are zero-mean unit-variance i.i.d. random variables with the finite eighth order moment. Then, as $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha$,

$$\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1 \xrightarrow{e.a.s.} 0 \tag{6.26}$$

where

$$\begin{aligned} \mathbf{B} &= \sum_{k=2}^{K} \mathbf{H}_k \mathbf{c}_k \mathbf{c}_k^H \mathbf{H}_k^H + \sigma^2 \mathbf{I} = \sum_{k=2}^{K} \mathbf{C}_k \mathbf{h}_k \mathbf{h}_k^H \mathbf{C}_k^H + \sigma^2 \mathbf{I} \\ \mathbf{C}_k &= \underline{\boldsymbol{\mathcal{C}}}_k + \tilde{\underline{\boldsymbol{\mathcal{C}}}}_k \\ \mathbf{H}_k &= \boldsymbol{\mathcal{H}}_k + \tilde{\boldsymbol{\mathcal{H}}}_k. \end{aligned} \tag{6.27}$$

Proof: The idea of the proof follows that of one of the results[3] in [14]. Note that

$$\|\underline{\boldsymbol{\mathcal{C}}}_1\|^2 \leq \|\underline{\boldsymbol{\mathcal{C}}}_1\|_F^2 \leq \|\mathbf{C}_1\|_F^2 = L \sum_{i=0}^{L_c-1} |c_1[i]|^2 \xrightarrow{a.s.} L \tag{6.28}$$

where the convergence in (6.28) is the direct consequence of the strong law of large numbers. It also follows from (6.28) that $\|\mathbf{C}_1\|^2 \leq L$. From (6.27) we have

$$\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 = \underline{\boldsymbol{\mathcal{C}}}_1^H \left(\mathbf{B} + \mathbf{E} + \underline{\tilde{\mathbf{w}}}_1 \underline{\tilde{\mathbf{w}}}_1^H\right)^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \tag{6.29}$$

where

$$\mathbf{E} = -\sum_{k=2}^{K} \boldsymbol{\mathcal{H}}_k \mathbf{c}_k \mathbf{c}_k \tilde{\boldsymbol{\mathcal{H}}}_k^H - \sum_{k=2}^{K} \tilde{\boldsymbol{\mathcal{H}}}_k \mathbf{c}_k \mathbf{c}_k \boldsymbol{\mathcal{H}}_k^H. \tag{6.30}$$

First of all, let us show that

$$\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \underline{\boldsymbol{\mathcal{C}}}_1^H \left(\mathbf{B} + \mathbf{E}\right)^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \xrightarrow{e.a.s.} 0. \tag{6.31}$$

Let us use the following inequality

$$\begin{aligned} &\left| \left[\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \underline{\boldsymbol{\mathcal{C}}}_1^H \left(\mathbf{B} + \mathbf{E}\right)^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \right]_{il} \right| \\ &\qquad = \frac{1}{1 + \underline{\tilde{\mathbf{w}}}_1^H (\mathbf{B} + \mathbf{E})^{-1} \underline{\tilde{\mathbf{w}}}_1} \left| \left[\underline{\boldsymbol{\mathcal{C}}}_1^H (\mathbf{B} + \mathbf{E})^{-1} \underline{\tilde{\mathbf{w}}}_1 \underline{\tilde{\mathbf{w}}}_1^H (\mathbf{B} + \mathbf{E})^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \right]_{il} \right| \\ &\qquad \leq \|\underline{\boldsymbol{\mathcal{C}}}_1\|^2 \|(\mathbf{B} + \mathbf{E})^{-1}\|^2 \|\underline{\tilde{\mathbf{w}}}_1\|^2 \end{aligned} \tag{6.32}$$

which follows from (6.29) and the matrix inversion lemma. Using (6.32) along with assumptions **A4** and **A5**, and the results of [14, Appendix I], it can be readily shown that

$$\|\underline{\tilde{\mathbf{w}}}_1\| \xrightarrow{a.s.} 0. \tag{6.33}$$

From the latter observation along with the facts that $\|\underline{\boldsymbol{\mathcal{C}}}_1\|$ and $\|(\mathbf{B} + \mathbf{E})^{-1}\|$ are both bounded,[4] we obtain (6.31).

Next, let us prove that

$$\underline{\boldsymbol{\mathcal{C}}}_1^H \left(\mathbf{B} + \mathbf{E}\right)^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \mathbf{C}_1^H \left(\mathbf{B} + \mathbf{E}\right)^{-1} \mathbf{C}_1 \xrightarrow{e.a.s.} 0. \tag{6.34}$$

[3]Note that this result of [14] applies to the difference between two random variables rather than two random matrices in the case when the spreading codes are columns of the Haar distributed matrix.

[4]The latter fact follows from $\|(\mathbf{B} + \mathbf{E})^{-1}\| \leq 1/\sigma^2$.

From (6.27) we have

$$\begin{aligned}\left|\left[\underline{\boldsymbol{\mathcal{C}}}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\underline{\boldsymbol{\mathcal{C}}}_1-\mathbf{C}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\mathbf{C}_1\right]_{il}\right| \\ \leq\left|\left[\mathbf{C}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1\right]_{il}\right|+\left|\left[\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\mathbf{C}_1\right]_{il}\right| \\ +\left|\left[\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1\right]_{il}\right|. \end{aligned} \tag{6.35}$$

For the first term in the right hand side of (6.35), we have

$$\left|\left[\mathbf{C}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1\right]_{il}\right|\leq\|\mathbf{C}_1\|\left\|(\mathbf{B}+\mathbf{E})^{-1}\right\|\left\|[\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1]_{\bullet l}\right\|. \tag{6.36}$$

Using Minkowski's inequality [14, Appendix I], it can be readily shown that

$$\left\|[\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1]_{\bullet l}\right\|\xrightarrow{a.s.}0. \tag{6.37}$$

Therefore, due to the boundedness of the other terms in the right hand side of (6.36), we have

$$\left|\left[\mathbf{C}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\tilde{\underline{\boldsymbol{\mathcal{C}}}}_1\right]_{il}\right|\xrightarrow{a.s.}0. \tag{6.38}$$

Similar approach can be used to show that the other two terms in the right hand side of (6.35) also converge almost surely to zero. This directly yields (6.34).

Now, let us show that

$$\mathbf{C}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\mathbf{C}_1-\mathbf{C}_1^H\mathbf{B}^{-1}\mathbf{C}_1\xrightarrow{e.a.s.}0. \tag{6.39}$$

Note that

$$[\mathbf{C}_1]_{\bullet l}=\mathbf{M}_{l-1}[\mathbf{C}_1]_{\bullet 1} \tag{6.40}$$

where $\mathbf{M}_{l-1}$ is the circulant matrix whose first column is $\mathbf{e}_l$. This matrix has the following properties:

$$\mathbf{M}_{-i}=\mathbf{M}_i^H=\mathbf{M}_i^{-1},\qquad\mathbf{M}_i\mathbf{M}_l=\mathbf{M}_{(i+l)\mathrm{mod}L_c}. \tag{6.41}$$

Using (6.40), we obtain

$$\begin{aligned}\left[\mathbf{C}_1^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\mathbf{C}_1\right]_{il}&=[\mathbf{C}_1^H]_{i\bullet}\left(\mathbf{B}+\mathbf{E}\right)^{-1}[\mathbf{C}_1]_{\bullet l} \\ &=\mathbf{c}_1^H\left(\mathbf{M}_{i-1}^H\left(\mathbf{B}+\mathbf{E}\right)^{-1}\mathbf{M}_{l-1}\right)\mathbf{c}_1. \end{aligned} \tag{6.42}$$

From [17, Lemma 1], it follows that

$$\mathbf{c}_1^H \left(\mathbf{M}_{i-1}^H \left(\mathbf{B} + \mathbf{E} \right)^{-1} \mathbf{M}_{l-1} \right) \mathbf{c}_1 - \frac{1}{L_c} \mathrm{tr} \left(\mathbf{M}_{i-1}^H \left(\mathbf{B} + \mathbf{E} \right)^{-1} \mathbf{M}_{l-1} \right) \xrightarrow{a.s.} 0. \quad (6.43)$$

Using (6.41)-(6.43), we obtain

$$\left[\mathbf{C}_1^H \left(\mathbf{B} + \mathbf{E} \right)^{-1} \mathbf{C}_1 \right]_{il} - \frac{1}{L_c} \mathrm{tr} \left(\mathbf{M}_{l-i} \left(\mathbf{B} + \mathbf{E} \right)^{-1} \right) \xrightarrow{a.s.} 0. \quad (6.44)$$

Similarly,

$$\left[\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1 \right]_{il} - \frac{1}{L_c} \mathrm{tr} \left(\mathbf{M}_{l-i} \mathbf{B}^{-1} \right) \xrightarrow{a.s.} 0. \quad (6.45)$$

According to (6.44) and (6.45), property (6.39) holds true if

$$\frac{1}{L_c} \mathrm{tr} \left(\mathbf{M}_{m-1} \left(\mathbf{B} + \mathbf{E} \right)^{-1} \right) - \frac{1}{L_c} \mathrm{tr} \left(\mathbf{M}_{m-1} \mathbf{B}^{-1} \right) \xrightarrow{a.s.} 0 \,, \quad m = 1, \ldots, L. \quad (6.46)$$

As only the first $L-1$ entries of $\underline{\tilde{\mathbf{w}}}_k = \tilde{\boldsymbol{\mathcal{H}}}_k \mathbf{c}_k$ are nonzero, we have that

$$\mathtt{rank}(\mathbf{E}) \leq 2L - 2. \quad (6.47)$$

Using (6.47), (6.46) can be proved [14, Appendix I], and, therefore, (6.39) is proved. Equation (6.26) directly follows from (6.31), (6.34), and (6.39). □

From (1.12), (1.15), and (6.27), it can be observed that each column of $\mathbf{C}_k$ is the circularly delayed version of the previous one and $\mathbf{H}_k$ is a circulant matrix. It is noteworthy that if the spreading codes $\mathbf{c}_k$ are augmented by the cyclic prefixes of length $L-1$, then, after removing the cyclic prefixes at the receiver, the received data covariance matrix is given by $\underline{\mathbf{R}} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{c}_k \mathbf{c}_k^H \mathbf{H}_k^H + \sigma^2 \mathbf{I}$ for any bounded L_c and K [14]. In such a case, (6.8), (6.16), and (6.21) should be modified by substituting $\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1$ in lieu of $\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1$. However, Proposition 1 shows that for our asymptotic case, such cyclic prefix augmentation is not required and the properties of $\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1$ and $\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1$ are identical. Therefore, hereafter we study the asymptotic properties of the latter matrix.

To justify the reliability of the blind MOE receivers, let us make a preliminary comparison of the asymptotic SINRs of the MMSE receiver (6.7) (that benefits from the exact knowledge of the channel vector $\mathbf{h}_1$) and the receivers (6.19) and (6.23). For the sake of

simplicity, we temporarily assume that the columns of $\mathbf{C}_k$ are statistically independent from each other. It is noteworthy that this assumption has been successfully used in [19] and [59] to obtain the asymptotic properties of the MMSE receiver. Moreover, as it will be shown in Section 6.4, the results derived using such a simplifying assumption offer reliable approximations for the scenarios in which $h_{k,l}$'s are zero-mean independent circular random variables.

Theorem 1: Let $\mathbf{c}_k$ be defined in Proposition 1 and the columns of $\mathbf{C}_k$ be statistically independent. Then, as $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha$,

$$\left(\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1\right)^{-1} - \frac{1}{\eta}\mathbf{I} \xrightarrow{e.a.s.} \mathbf{0} \tag{6.48}$$

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) - \|\mathbf{h}_1\|^2 \eta \xrightarrow{a.s.} 0 \tag{6.49}$$

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\|\bar{\mathbf{h}}\|^2}\eta \xrightarrow{a.s.} 0 \tag{6.50}$$

$$\mathbf{h}_{\mathrm{C}} - \xi \mathbf{h}_1 \xrightarrow{e.a.s.} \mathbf{0} \tag{6.51}$$

where $\eta = \frac{1}{L_c}\mathrm{tr}(\mathbf{B}^{-1})$ and ξ is a scaling factor. Moreover, if the empirical distribution of the received user powers $\{\|\mathbf{h}_2\|^2, \|\mathbf{h}_3\|^2, \ldots, \|\mathbf{h}_K\|^2\}$ converges almost surely to a nonrandom distribution $F(p)$, then

$$\eta \xrightarrow{a.s.} \beta^* \tag{6.52}$$

where β^* is the unique deterministic positive solution to

$$\beta = \left(\sigma^2 + \alpha \int_0^\infty \frac{p}{1+p\beta}\, dF(p)\right)^{-1}. \tag{6.53}$$

Proof: First of all, let us show that $\lambda_{\max}(\mathbf{B})$ is upper-bounded. From (6.27) we have

$$\mathbf{B} = \mathbf{C}_{\bar{1}} \mathbf{P}_{\bar{1}} \mathbf{C}_{\bar{1}}^H + \sigma^2 \mathbf{I} \tag{6.54}$$

where $\mathbf{C}_{\bar{1}} = [\mathbf{C}_2, \ldots, \mathbf{C}_K]$ and $\mathbf{P}_{\bar{1}} = \mathrm{diag}\{\mathbf{h}_2\mathbf{h}_2^H, \ldots, \mathbf{h}_K\mathbf{h}_K^H\}$. From (6.54), it follows that

$$\lambda_{\max}(\mathbf{B}) \leq \sigma^2 + \lambda_{\max}(\mathbf{C}_{\bar{1}}\mathbf{C}_{\bar{1}}^H)\lambda_{\max}(\mathbf{P}_{\bar{1}}). \tag{6.55}$$

Also, using [5, Theorem 2], we have

$$\lambda_{\max}(\mathbf{C}_{\bar{1}}\mathbf{C}_{\bar{1}}^H) \xrightarrow{a.s.} (1+\sqrt{\alpha L})^2 \tag{6.56}$$

and from assumption **A4** it follows that

$$\lambda_{\max}(\mathbf{P}_{\bar{1}}) < S. \tag{6.57}$$

Using (6.56) and (6.57) in (6.55), we have that $\lambda_{\max}(\mathbf{B})$ is upper-bounded with probability one.

According to (6.54), we have

$$\lambda_{\max}(\mathbf{B}^{-1}) \leq 1/\sigma^2. \tag{6.58}$$

As the columns of $\mathbf{C}_1$ are statistically independent from each other and $\mathbf{B}^{-1}$, it directly follows from (6.58) and [19, Corollary 1] that

$$\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1 - \eta \mathbf{I} \xrightarrow{e.a.s.} \mathbf{0}. \tag{6.59}$$

Also, note that

$$\eta = \frac{1}{L_c}\mathrm{tr}(\mathbf{B}^{-1}) \geq \lambda_{\min}(\mathbf{B}^{-1}) = \frac{1}{\lambda_{\max}(\mathbf{B})}. \tag{6.60}$$

As $\lambda_{\max}(\mathbf{B})$ is almost surely upper-bounded, it follows from (6.60) that $\eta > 0$ with probability one. The latter observation and (6.59) directly yield (6.48), while (6.8) along with (6.59) prove (6.49) [19]. Furthermore, (6.50) can be directly obtained from (6.21), (6.26) and (6.48). To prove (6.51), we first note that from Wielandt-Hoffman theorem [95, Chapter 2]) along with (6.15) and (6.48), it follows that

$$\nu - \left(\frac{1}{\eta} + \|\mathbf{h}_1\|^2\right) \xrightarrow{a.s.} 0. \tag{6.61}$$

Rewriting (6.16) in the following equivalent form

$$\left((\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} + \mathbf{h}_1 \mathbf{h}_1^H\right) \mathbf{h}_{\mathrm{C}} - \nu \mathbf{h}_{\mathrm{C}} = 0 \tag{6.62}$$

and using Proposition 1 along with (6.48) and (6.61), we have that, as $L_c \to \infty$ with $K/L_c \to \alpha$,

$$\left(\frac{1}{\eta}\mathbf{I} + \mathbf{h}_1 \mathbf{h}_1^H\right) \mathbf{h}_{\mathrm{C}} - \left(\frac{1}{\eta} + \|\mathbf{h}_1\|^2\right) \mathbf{h}_{\mathrm{C}} \xrightarrow{e.a.s.} 0. \tag{6.63}$$

Equation (6.51) directly follows from (6.63). Proof of (6.52) is given in [19, Theorem 7]. □

It can be seen from Theorem 1 that if (6.52) holds true, then both $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$ converge almost surely to *deterministic* values.[5] It also directly follows from (6.49)-(6.51) that

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})) - \mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) \xrightarrow{a.s.} 0. \tag{6.64}$$

Equation (6.64) shows that, as L_c and K go to infinity, the SINR of the *blind* MOE receiver (6.23) approaches $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ (which is the maximum possible value of SINR of any linear receiver). Note that the simplifying assumption that the columns of $\mathbf{C}_k$ are statistically independent is essential to obtain the results of Theorem 1 and if we drop this assumption, (6.48), (6.49)-(6.52), and (6.64) do not hold true in general. For instance, as it will be shown in Section 6.5, (6.50) does not accurately predict $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$ in the downlink transmission case. This fact motivates an asymptotic performance analysis of the MOE receiver (6.19) without using the aforementioned assumption of statistical independence of the columns of $\mathbf{C}_k$. The rest of this chapter is devoted to such analysis. To simplify the subsequent derivations up to a reasonable level, we confine our analysis to the circular Gaussian spreading codes, that is, we assume that:

A5: Entries of $\mathbf{c}_k$ are zero-mean i.i.d. circular Gaussian random variables of variance $1/L_c$.

Note that **A5** has been frequently used in the literature to analyze the asymptotic performance of the MMSE receiver [14], [19], and to derive reduced-rank MMSE multiuser techniques [28].

The following theorem is the key result to our subsequent analysis.

Theorem 2: If L_c and K converge to $+\infty$ with $\frac{K}{L_c} \to \alpha$, then

$$\left(\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1\right)^{-1} - \mathbf{T}^{-1} \xrightarrow{e.i.p.} \mathbf{0} \tag{6.65}$$

where $\mathbf{T}$ is a positive definite Hermitian Toeplitz matrix with

$$[\mathbf{T}]_{1m} = \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{e^{-j\frac{2\pi}{L_c}(m-1)(l-1)}}{\phi_l}, \qquad m = 1, \ldots, L \tag{6.66}$$

[5]The fact that $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ converges to $\|\mathbf{h}_1\|^2 \beta^*$ has been originally proved in [79] and [19].

and ϕ_l is the unique real and positive solution to

$$\phi_l = \sigma^2 + \frac{1}{L_c}\sum_{k=2}^{K} \frac{\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2}{1+\frac{1}{L_c}\sum_{n=1}^{L_c}\left|h_k(e^{j\frac{2\pi}{L_c}(n-1)})\right|^2/\phi_n}, \qquad l=1,\ldots,L_c. \tag{6.67}$$

Proof: As $\mathbf{H}_k$ and $\mathbf{M}_m$ are circulant matrices, we have

$$\mathbf{H}_k = \mathbf{F}\mathbf{\Lambda}_k\mathbf{F}^H, \qquad k=1,\ldots,K \tag{6.68}$$

$$\mathbf{M}_{m-1} = \mathbf{F}\mathbf{D}_{m-1}\mathbf{F}^H, \qquad m=1,\ldots,L \tag{6.69}$$

where $\mathbf{F}$ is an $L_c \times L_c$ matrix with

$$[\mathbf{F}]_{st} = \frac{1}{\sqrt{L_c}} e^{j\frac{2\pi}{L_c}(s-1)(t-1)} \tag{6.70}$$

and

$$\mathbf{\Lambda}_k = \operatorname{diag}\left(\sqrt{L_c}\mathbf{F}^H[\mathbf{H}_k]_{\bullet 1}\right) \tag{6.71}$$

$$\mathbf{D}_{m-1} = \operatorname{diag}\left(\sqrt{L_c}\mathbf{F}^H\mathbf{e}_m\right). \tag{6.72}$$

From (6.24) and (6.71), it follows that

$$[\mathbf{\Lambda}_k]_{nn} = h_k\big(e^{j\frac{2\pi}{L_c}(n-1)}\big). \tag{6.73}$$

The following result obtained from Theorems 7.1 and 7.2 of [26, Chapter 7] (also see [14, Eq. (60)]) is essential for our proof of Theorem 2.

Theorem 3: Let $\mathbf{\Xi}$ be an $L_c \times K$ random matrix with zero-mean entries and *independent columns*. Let $L_c, K \to +\infty$ with $0 < \lim_{L_c \longrightarrow +\infty}(K/L_c) = \alpha < \infty$ and

$$\sup_{L_c} \max_{\substack{n=1,\ldots,L_c \\ k=1,\ldots,K}} \left\{ \sum_{n=1}^{L_c} \zeta_{nk}^2 + \sum_{k=1}^{K} \zeta_{nk}^2 \right\} < \infty \tag{6.74}$$

where $\mathrm{E}\{|[\mathbf{\Xi}]_{nk}|^2\} = \zeta_{nk}^2$. Moreover, let for every $\tau > 0$

$$f(\tau) = \lim_{L_c\to\infty} \max_{\substack{n=1,\ldots,L_c \\ k=1,\ldots,K}} \Big\{ \sum_{n=1}^{L_c} \mathrm{E}\left\{|[\mathbf{\Xi}]_{nk}|^2 \mathrm{I}\left\{|[\mathbf{\Xi}]_{nk}| > \tau\right\}\right\} + \sum_{k=1}^{K} \mathrm{E}\left\{|[\mathbf{\Xi}]_{nk}|^2 \mathrm{I}\left\{|[\mathbf{\Xi}]_{nk}| > \tau\right\}\right\} \Big\} = 0. \tag{6.75}$$

Then, as $L_c \to \infty$,

$$\left[\left(\mathbf{\Xi}\mathbf{\Xi}^H - z\mathbf{I}\right)^{-1}\right]_{ll} - 1/\phi_l(z) \xrightarrow{i.p.} 0 \tag{6.76}$$

where

$$\phi_l(z) = -z + \sum_{k=1}^{K} \frac{\zeta_{lk}^2}{1 + \sum_{n=1}^{L_c} \frac{\zeta_{nk}^2}{\phi_n(z)}}, \qquad l = 1, \ldots, L_c. \tag{6.77}$$

The equation system in (6.77) has a unique set of solutions which are analytical in

$$\mathcal{K} = \{z \,|\, \mathrm{Im}(z) > 0,\, \mathrm{Im}(\phi_l(z)) < 0,\, l = 1, \ldots, L_c\}. \tag{6.78}$$

We will also need the following lemma.

Lemma 1: Conditions (6.74) and (6.75) hold true for

$$\mathbf{\Xi}_c = [\mathbf{\Lambda}_2\mathbf{c}_2 \;\; \mathbf{\Lambda}_3\mathbf{c}_3 \;\; \cdots \;\; \mathbf{\Lambda}_K\mathbf{c}_K]. \tag{6.79}$$

Proof: First, note that $[\mathbf{\Xi}_c]_{nk} \sim \mathcal{NC}(0, \zeta_{nk}^2)$ where

$$\zeta_{nk} = \left|h_{k+1}\left(e^{j\frac{2\pi}{L_c}(n-1)}\right)\right|/\sqrt{L_c}. \tag{6.80}$$

Inequality (6.74) directly follows from (6.25) and (6.80). To prove (6.75), let

$$|\theta_{nk}|^2 = 2|[\mathbf{\Xi}_c]_{nk}|^2/\zeta_{nk}^2. \tag{6.81}$$

It can be observed that $|\theta_{nk}|^2$ is χ^2-distributed with $r = 2$ degrees of freedom. Hence,

$$\begin{aligned}
\mathrm{E}\left\{|[\mathbf{\Xi}_c]_{nk}|^2\,\mathrm{I}\left\{|[\mathbf{\Xi}_c]_{nk}| > \tau\right\}\right\} &= \frac{\zeta_{nk}^2}{2}\mathrm{E}\left\{|\theta_{nk}|^2\,\mathrm{I}\left\{|\theta_{nk}| > \frac{\sqrt{2}\tau}{\zeta_{nk}}\right\}\right\} \\
&= \frac{\zeta_{nk}^2}{4}\int_{2\tau^2/\zeta_{nk}^2}^{\infty} x e^{-x/2} dx \\
&= \tau^2 e^{-\tau^2/\zeta_{nk}^2} + \zeta_{nk}^2 e^{-\tau^2/\zeta_{nk}^2} \\
&\leq \tau^2 e^{-L_c\tau^2/B} + (B/L_c)\, e^{-L_c\tau^2/B}.
\end{aligned} \tag{6.82}$$

From (6.75) and (6.82), it follows that

$$f(\tau) \leq \lim_{L_c\to\infty} L_c(1+\alpha)\left(\tau^2 e^{-L_c\tau^2/B} + (B/L_c)\, e^{-L_c\tau^2/B}\right) = 0. \tag{6.83}$$

This completes the proof of Lemma 1.

Let us now prove that

$$\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1 - \mathbf{T} \xrightarrow{e.i.p.} 0. \tag{6.84}$$

From (6.45), we have that

$$\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1 - \mathbf{X} \xrightarrow{e.a.s.} 0 \tag{6.85}$$

where $\mathbf{X}$ is a Hermitian Toeplitz matrix with

$$[\mathbf{X}]_{1\bullet} = \left[\frac{1}{L_c} \mathrm{tr}\left(\mathbf{M}_0 \mathbf{B}^{-1}\right), \frac{1}{L_c} \mathrm{tr}\left(\mathbf{M}_1 \mathbf{B}^{-1}\right), \ldots, \frac{1}{L_c} \mathrm{tr}\left(\mathbf{M}_{L-1} \mathbf{B}^{-1}\right) \right]. \tag{6.86}$$

Hence, it is sufficient to show that $\mathbf{X} - \mathbf{T} \xrightarrow{e.i.p.} 0$, or, equivalently

$$\frac{1}{L_c} \mathrm{tr}\left(\mathbf{M}_{m-1} \mathbf{B}^{-1}\right) - [\mathbf{T}]_{1m} \xrightarrow{i.p.} 0\,, \qquad m = 1, \ldots, L. \tag{6.87}$$

Using (6.27) along with (6.68) and (6.69), after straightforward manipulation it can be shown that

$$\frac{1}{L_c} \mathrm{tr}\left(\mathbf{M}_{m-1} \mathbf{B}^{-1}\right) = \frac{1}{L_c} \sum_{l=1}^{L_c} e^{-j\frac{2\pi}{L_c}(m-1)(l-1)} \left[\left(\sum_{k=2}^{K} \mathbf{\Lambda}_k \mathbf{F}^H \mathbf{c}_k \mathbf{c}_k^H \mathbf{F} \mathbf{\Lambda}_k^H + \sigma^2 \mathbf{I} \right)^{-1} \right]_{l,l}. \tag{6.88}$$

As $\mathbf{c}_k, k = 2, \ldots, K$ are Gaussian, their distribution remains invariant to left multiplication by any deterministic unitary matrix. Using the latter property, we obtain that for showing (6.87), we have to prove that

$$\frac{1}{L_c} \sum_{l=1}^{L_c} e^{-j\frac{2\pi}{L_c}(m-1)(l-1)} \left[\left(\mathbf{\Xi}_c \mathbf{\Xi}_c^H + \sigma^2 \mathbf{I} \right)^{-1} \right]_{l,l} - [\mathbf{T}]_{1m} \xrightarrow{i.p.} 0. \tag{6.89}$$

As $\mathbf{\Xi}_c$ is a zero-mean random matrix with independent columns that satisfies (6.74) and (6.75), it follows from Theorem 3 that

$$\left[\left(\mathbf{\Xi}_c \mathbf{\Xi}_c^H + \sigma^2 \mathbf{I} \right)^{-1} \right]_{l,l} - \frac{1}{\phi_l} \xrightarrow{i.p.} 0\,, \qquad l = 1, \ldots, L_c \tag{6.90}$$

where ϕ_l is given by (6.67). Note that (6.67) is obtained by substituting

$$\zeta_{nk} = \left| h_{k+1}\left(e^{j\frac{2\pi}{L_c}(n-1)}\right) \right| / \sqrt{L_c} \tag{6.91}$$

in (6.77) and computing the result for $z \to -\sigma^2$. Then, (6.89) directly follows from the definition of convergence in probability and the fact that (6.90) holds for all $l = 1, \ldots, L_c$. According to Theorem 4, the set of solutions to (6.67) is unique. We also remark that

$$\sigma^2 \leq \phi_l \leq \sigma^2 + \alpha B\,, \qquad l = 1, \ldots, L_c \tag{6.92}$$

where the left hand side inequality follows from

$$\left[\left(\boldsymbol{\Xi}_c \boldsymbol{\Xi}_c^H + \sigma^2 \mathbf{I}\right)^{-1}\right]_{l,l} \leq 1/\sigma^2\,, \qquad l = 1, \ldots, L_c \tag{6.93}$$

and the right hand side inequality can be easily obtained from (6.67). As ϕ_l for $l = 1, \ldots, L_c$ are bounded positive scalars, it can be shown that the Hermitian Toeplitz matrix $\mathbf{T}$ is positive definite [108, Lemma 1], and, therefore invertible. Equation (6.65) directly follows from the latter fact and (6.84). □

Note that Theorem 2 along with (6.8) and (6.21) yields

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) - \mathbf{h}_1^H \mathbf{T} \mathbf{h}_1 \xrightarrow{i.p.} 0 \tag{6.94}$$

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\bar{\mathbf{h}}^H \mathbf{T}^{-1} \bar{\mathbf{h}}} \xrightarrow{i.p.} 0. \tag{6.95}$$

Interestingly, an asymptotic expression similar to (6.94) has been earlier obtained in [14]. In contrast to Theorem 1 that assumes the columns of $\mathbf{C}_k$ to be statistically independent to show that $\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1$ converges to a scaled version of the identity matrix, Theorem 2 proves that $\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1$ in fact converges to the Hermitian Toeplitz matrix $\mathbf{T}$ whose off-diagonal entries, in general, are nonzero. Note that, due to the nonzero off-diagonal entries of $\mathbf{T}$, matrix $\mathbf{T}^{-1}$ deviates from a scaled version of the identity matrix. As can be observed from (6.16), it follows from the latter fact that, in general, $\mathbf{h}_{\mathrm{C}}$ does not asymptotically converge to a scaled version of $\mathbf{h}_1$, and, consequently, $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ does not reach the value of $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$.

The above discussion motivates our especial interest to the particular cases wherein $\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1$ converges to a scaled version of the identity matrix, that is $[\mathbf{T}]_{1m} = 0$ for $m \neq 1$. Let us now discuss a notable practical case which results in such a convergence.

Corollary 1: Assume that the users have flat fading channels, that is,

$$\mathbf{h}_k = h_{k,q_k}\mathbf{e}_{q_k}\,, \qquad k = 1,\ldots,K \tag{6.96}$$

where $q_k \in \{0,1,\ldots,L-1\}$. Then

$$\left(\mathbf{C}_1^H\mathbf{B}^{-1}\mathbf{C}_1\right)^{-1} - \phi^*\mathbf{I} \xrightarrow{e.i.p.} 0 \tag{6.97}$$

where ϕ^* is the unique positive solution to

$$\phi = \sigma^2 + \frac{1}{L_c}\sum_{k=2}^{K}\frac{|h_{k,q_k}|^2}{1+|h_{k,q_k}|^2/\phi}. \tag{6.98}$$

Moreover, if the empirical distribution of the user powers $\{|h_{2,q_2}|^2, |h_{3,q_3}|^2, \ldots, |h_{K,q_K}|^2\}$ converges almost surely to the nonrandom distribution $F(p)$, then

$$\phi^* \xrightarrow{a.s.} 1/\beta^*. \tag{6.99}$$

Proof: From (6.96), it follows that for $l = 1,\ldots,L_c$ and $k = 1,\ldots,K$,

$$\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2 = |h_{k,q_k}|^2. \tag{6.100}$$

Inserting (6.100) into (6.67), we obtain

$$\phi_l = \sigma^2 + \frac{1}{L_c}\sum_{k=2}^{K}\frac{|h_{k,q_k}|^2}{1+\dfrac{1}{L_c}\displaystyle\sum_{n=1}^{L_c}\frac{|h_{k,q_k}|^2}{\phi_n}}\,, \qquad l = 1,\ldots,L_c. \tag{6.101}$$

As the right-hand side of (6.101) is independent from l, we have that $\phi_1 = \cdots = \phi_{L_c} = \phi^*$ where ϕ^* is the unique positive solution to (6.98). The latter result along with (6.66) yields

$$[\mathbf{T}]_{1m} = \frac{1}{L_c\phi^*}\sum_{l=1}^{L_c} e^{-j\frac{2\pi}{L_c}(m-1)(l-1)} = \delta_{1m}/\phi^*\,, \qquad m = 1,\ldots,L \tag{6.102}$$

and (6.97) immediately follows from (6.102). To prove (6.99), we first consider the function

$$f(x) = \frac{\sigma^2}{x} + \alpha\int_0^\infty \frac{p/x}{1+p/x}\,dF(p)\,, \quad x > 0. \tag{6.103}$$

Note that $f(x)$ is continuous and injective. As β^* is the solution to (6.53), we have [79]

$$f\left(1/\beta^*\right) = 1. \tag{6.104}$$

Let $F^{(L_c)}(p)$ denote the empirical distribution of $\{|h_{2,q_2}|^2, |h_{3,q_3}|^2, \ldots, |h_{K,q_K}|^2\}$, that is,

$$F^{(L_c)}(p) = \frac{1}{K-1} \sum_{k=2}^{K} \mathrm{I}\left\{|h_{k,q_k}|^2 \leq p\right\}. \tag{6.105}$$

As ϕ^* is the solution to (6.98), it immediately follows that

$$\phi^* - \left(\sigma^2 + \frac{K-1}{L_c} \int_0^\infty \frac{p}{1+p/\phi^*}\, dF^{(L_c)}(p)\right) = 0. \tag{6.106}$$

From (6.92) we have that $\phi^* \geq \sigma^2$, and, hence, the integrand in (6.106) is a bounded continuous function of p. Using this property along with the fact that the empirical distribution function $F^{(L_c)}(p)$ almost surely converges at $L_c \to \infty$ to the deterministic function $F(p)$, we obtain

$$\phi^* - \left(\sigma^2 + \alpha \int_0^\infty \frac{p}{1+p/\phi^*}\, dF(p)\right) \xrightarrow{a.s.} 0. \tag{6.107}$$

Dividing both sides of (6.107) by ϕ^* and using (6.104), we obtain

$$f(\phi^*) \xrightarrow{a.s.} f\left(1/\beta^*\right). \tag{6.108}$$

As $f(\cdot)$ is injective and continuous in the neighborhood of $1/\beta^*$, (6.99) directly follows from (6.108). This completes the proof. □

It follows from Corollary 1 that if (6.96) holds, then, using (6.97) in (6.8) and (6.21), we have

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) - |h_{1,q_1}|^2/\phi^* \xrightarrow{i.p.} 0 \tag{6.109}$$

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \frac{|[\bar{\mathbf{h}}]_{q_1}|^2 |h_{1,q_1}|^2}{\|\bar{\mathbf{h}}\|^2 \phi^*} \xrightarrow{i.p.} 0. \tag{6.110}$$

It should be mentioned that, using a different approach, (6.109) has been obtained for the flat fading channels in [79]. Note also that, following similar steps as in the proof of (6.51), it can be shown from (6.97) that $\mathbf{h}_{\mathrm{C}} - \xi \mathbf{e}_{q_1} \xrightarrow{e.i.p.} 0$ for some scaling factor ξ. From the latter result and (6.109)-(6.110) it immediately follows that $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})) - \mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) \xrightarrow{i.p.} 0$. Two interesting particular cases are worth discussing:

- Consider the flat fading case and assume that the receiver is not perfectly synchronized with the user-of-interest, that is, $\mathbf{h}_1 = h\mathbf{e}_{q_1}$ where h is a scalar and q_1 is an unknown integer between 1 and L. Then, as $\mathbf{h}_{\mathrm{C}}$ converges to a scaled version of $\mathbf{e}_{q_1}$, the Capon channel estimation technique can be applied to retrieve q_1 and obtain the exact timing of the user-of-interest.

- Let the user channels be flat fading but at the receiver it is mistakenly assumed that the unknown user-of-interest channel is frequency-selective and is modelled as an unknown vector of length L. Let the Capon receiver (6.23) be obtained based on this erroneous channel model. Then, the SINR performances of such a Capon receiver and that of the optimum MMSE receiver (which uses the knowledge of the channel to achieve the maximum SINR) are asymptotically equal.

6.4 Channel distribution effect on the receiver performance

To study the effect of the user channels on the performances of the Capon channel estimation technique and the blind MOE receiver (6.19), one can use (6.24) to obtain [14]

$$\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2 = \sum_{s=-(L-1)}^{L-1} r_{k,s} e^{j\frac{2\pi}{L_c}(l-1)s} \tag{6.111}$$

where

$$r^*_{k,-s} = r_{k,s} = \sum_{l=s}^{L-1} h^*_{k,l} h_{k,l-s}\,, \qquad s = 0, \ldots, L-1. \tag{6.112}$$

Inserting (6.111) in (6.67), the latter equation can be rewritten as

$$\phi_l = \sigma^2 + \sum_{s=-(L-1)}^{L-1} e^{j\frac{2\pi}{L_c}(l-1)s} \left(\frac{1}{L_c}\sum_{k=2}^{K}\frac{r_{k,s}}{f_k}\right) \tag{6.113}$$

where

$$f_k = 1 + \sum_{p=-(L-1)}^{L-1} r_{k,p} \left(\frac{1}{L_c}\sum_{n=1}^{L_c}\frac{e^{j\frac{2\pi}{L_c}(n-1)p}}{\phi_n}\right). \tag{6.114}$$

Let us use a common for the uplink case assumption that the coefficients $h_{k,l}$ are zero-mean circular random variables that are independent but not necessarily identically distributed across k and l [14]. In such a case, it directly follows from (6.112) that

$$\mathrm{E}\{r_{k,s}\} = \mathrm{E}\left\{\|\mathbf{h}_k\|^2\right\}\delta_s\,, \quad \mathrm{E}\{r_{k,s}r^*_{k,p}\} = \xi_{k,s}\delta_{s,p} \tag{6.115}$$

where, according to **A4**, the real scalar $\xi_{k,s}$ can be shown to be upper-bounded. Using (6.114) and (6.115), it can be observed that f_k is a function of $2L-1$ random variables $r_{k,-(L-1)},\ldots,r_{k,L-1}$ among which only one is correlated with $r_{k,s}$. Then, it can be concluded that $1/f_k$ and $r_{k,s}$ are approximately uncorrelated with each other. The following theorem establishes several important properties resulting from this observation.

Theorem 4: Let $h_{k,l}$ be zero-mean circular random variables independent across k and l, and let $1/f_k$ and $r_{k,s}$ be uncorrelated. Then

$$\frac{1}{L_c}\sum_{k=2}^{K}\frac{r_{k,s}}{f_k} \xrightarrow{i.p.} 0\,, \qquad s=\pm1,\pm2,\ldots,\pm(L-1) \tag{6.116}$$

and

$$\left(\mathbf{C}_1^H\mathbf{B}^{-1}\mathbf{C}_1\right)^{-1} - \phi^*_{\mathbf{h}}\mathbf{I} \xrightarrow{e.i.p.} 0 \tag{6.117}$$

where $\phi^*_{\mathbf{h}}$ is the unique positive solution to

$$\phi_{\mathbf{h}} = \sigma^2 + \frac{1}{L_c}\sum_{k=2}^{K}\frac{\|\mathbf{h}_k\|^2}{1+\|\mathbf{h}_k\|^2/\phi_{\mathbf{h}}}. \tag{6.118}$$

Moreover, if the empirical distribution of the received user powers $\{\|\mathbf{h}_2\|^2,\|\mathbf{h}_3\|^3,\ldots,\|\mathbf{h}_K\|^2\}$ converges almost surely to the nonrandom distribution $F(p)$, then

$$\phi^*_{\mathbf{h}} \xrightarrow{a.s.} 1/\beta^*. \tag{6.119}$$

Proof: From (6.115) and the assumption that $1/f_k$ and $r_{k,s}$ are uncorrelated, it follows that

$$\mathrm{E}\left\{r_{k,s}/f_k\right\} = 0\,, \qquad s=\pm1,\pm2,\ldots,\pm(L-1). \tag{6.120}$$

Using (6.115) and the fact that $f_k > 1$ we also have that

$$\mathrm{E}\left\{\left|r_{k,s}/f_k\right|^2\right\} < \mathrm{E}\{|r_{k,s}|^2\} = \xi^2_{k,s} \tag{6.121}$$

and, therefore, $\mathrm{E}\left\{|r_{k,s}/f_k|^2\right\}$ is upper-bounded. As $\mathbf{h}_k$, and, therefore, $r_{k,s}/f_k$ are independent for different indices k, (6.116) directly follows from (6.120) and (6.121). Using (6.116) in (6.113), we obtain

$$\phi_l - \varphi \xrightarrow{i.p.} 0\,, \qquad l = 1, \ldots, L_c \tag{6.122}$$

where

$$\varphi = \sigma^2 + \frac{1}{L_c}\sum_{k=2}^{K} r_{k,0}/f_k. \tag{6.123}$$

From (6.122) and (6.66) it follows that the off-diagonal entries of $\mathbf{T}$ converge in probability to zero while $[\mathbf{T}]_{11} - 1/\varphi \xrightarrow{i.p.} 0$. Therefore, we have

$$\left(\mathbf{C}_1^H \mathbf{B}^{-1} \mathbf{C}_1\right)^{-1} - \varphi \mathbf{I} \xrightarrow{e.i.p.} 0. \tag{6.124}$$

Hence, to prove (6.117) it is sufficient to show that

$$\varphi - \phi_{\mathbf{h}}^* \xrightarrow{i.p.} 0. \tag{6.125}$$

To prove (6.125), using (6.122) in (6.114) we obtain

$$f_k - \left(1 + \sum_{p=-(L-1)}^{L-1} \frac{r_{k,p}}{\varphi}\left(\frac{1}{L_c}\sum_{n=1}^{L_c} e^{j\frac{2\pi}{L_c}(n-1)p}\right)\right) \xrightarrow{i.p.} 0 \tag{6.126}$$

or, equivalently,

$$f_k - (1 + r_{k,0}/\varphi) \xrightarrow{i.p.} 0. \tag{6.127}$$

From (6.123), (6.127) and the fact that $r_{k,0} = \|\mathbf{h}_k\|^2$, it follows that

$$\varphi - \left(\sigma^2 + \frac{1}{L_c}\sum_{k=2}^{K} \frac{\|\mathbf{h}_k\|^2}{1 + \|\mathbf{h}_k\|^2/\varphi}\right) \xrightarrow{i.p.} 0. \tag{6.128}$$

Introducing a continuous and injective function

$$g(x) = \frac{\sigma^2}{x} + \frac{1}{L_c}\sum_{k=2}^{K} \frac{\|\mathbf{h}_k\|^2}{x + \|\mathbf{h}_k\|^2} \tag{6.129}$$

for $x > 0$, it can be easily observed from (6.118) and (6.128) that

$$g(\varphi) - g(\phi_{\mathbf{h}}^*) \xrightarrow{i.p.} 0 \tag{6.130}$$

which directly proves (6.125). The proof of (6.119) is similar to that of (6.99) and, therefore, is omitted. □

Note that if the assumptions of Theorem 4 hold, then using (6.117) and following the same steps as contained in the proof of (6.49)-(6.51), it can be readily shown that

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) - \|\mathbf{h}_1\|^2/\phi^*_{\mathbf{h}} \xrightarrow{i.p.} 0 \tag{6.131}$$

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - |\bar{\mathbf{h}}^H\mathbf{h}_1|^2/(\|\bar{\mathbf{h}}\|^2\phi^*_{\mathbf{h}}) \xrightarrow{i.p.} 0 \tag{6.132}$$

$$\mathbf{h}_{\mathrm{C}} - \xi\mathbf{h}_1 \xrightarrow{e.i.p.} \mathbf{0} \tag{6.133}$$

where ξ is a scaling factor.

From (6.131)-(6.133) it directly follows that

$$\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})) - \mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) \xrightarrow{i.p.} 0. \tag{6.134}$$

Ignoring the differences between the types of convergence, it can be concluded that (6.133) and (6.134) establish essentially the same results as (6.51) and (6.64), respectively. This verifies that if the assumptions of Theorem 4 hold, then (6.51) and (6.64) (which have been obtained under the assumption that the delayed versions of the spreading codes are statistically independent from each other) are still valid results. It is also worth noting from (6.52) and (6.119) that if the empirical distribution of the received user powers converges almost surely to the arbitrary deterministic distribution $F(p)$, then η and $1/\phi^*_{\mathbf{h}}$ both converge to the same deterministic value of β^*. In such a case, the asymptotic SINRs in (6.49) and (6.50) are also equal to those in (6.131) and (6.132), respectively. We emphasize that in practical uplink scenarios $1/f_k$ and $r_{k,s}$ have rather small correlation. Therefore, (6.116) and, consequently, (6.117) and (6.131)-(6.134) do not hold in the exact sense. However, Theorem 4 shows that if the channel coefficients are zero-mean circular and independent random variables, then the results of Theorem 1 are useful approximations.[6]

[6]It was earlier mentioned in [14] that if $h_{k,l}$ are independent across k and l, then $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ is approximately equal to the asymptotic expression in [19] derived using the statistical independence of the delayed user spreading codes. Theorem 4 provides theoretical support to the latter argument by rigorously proving equations (6.116)-(6.117) and (6.119)-(6.133) and deriving exact conditions under which these expressions hold true.

It is worth noting that the results of Theorems 1 and 3 are not necessarily relevant in the downlink case, where we have

$$\mathbf{h}_k = A_k \check{\mathbf{h}}. \qquad k = 1, \ldots, K \tag{6.135}$$

In (6.135), A_k is the amplitude allocated by the base station to the kth user, and $\check{\mathbf{h}}$ represents the physical normalized channel between the base station and the user-of-interest.

From (6.135) it can be observed that the statistical independence of $h_{k,l}$ across k and l is no longer a proper assumption and, therefore, the results of Theorems 1 and 3 are not relevant anymore. Indeed, using (6.135) in (6.112), we have $r_{k,s} = p_k r_s$ $(s = 0, \pm 1, \ldots, \pm(L-1))$ with $r_s = \sum_{l=s}^{L-1} h_{1,l}^* h_{1,l-s}$. Hence,

$$\frac{1}{L_c} \sum_{k=2}^{K} \frac{r_{k,s}}{f_k} = \frac{r_s}{L_c} \sum_{k=2}^{K} \frac{p_k}{f_k}, \qquad s = 0, \pm 1, \pm 2, \ldots, \pm(L-1). \tag{6.136}$$

As $p_k/f_k > 0$ for all k and r_s is a constant, unlike (6.116), Equation (6.136) does not converge to zero for any $s = \pm 1, \ldots, \pm(L-1)$. From the latter observation and (6.113) it follows that, in contrary to the uplink scenario, in the downlink case ϕ_l are not asymptotically equal to the common value of $\sigma^2 + \frac{1}{L_c} \sum_{k=2}^{K} \frac{r_{k,0}}{f_k}$ for $l = 1, \ldots, L_c$, and, therefore, the off-diagonal entries of $\mathbf{T}$ do not necessarily vanish to zero. This, in turn, implies that $\mathbf{h}_{\mathrm{C}}$ and $\mathbf{h}_1$ are different and that $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ is higher than $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$. The above discussion establishes the interesting result that if the Capon technique (6.14) is used to estimate the channel vector $\mathbf{h}_1$, then the estimation performance is expected to be better if $\mathbf{h}_1$ is the uplink rather than downlink channel. Also, $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ tends to be closer to $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ in the uplink rather than in the downlink case.

6.5 Simulations

To verify that our analysis is applicable to non-Gaussian spreading codes, throughout all numerical examples, the entries of the these codes have been randomly generated from $\pm 1/\sqrt{L_c}$. In Figs. 6.1-6.4, a total of 200 sets of spreading codes of lengths $L_c = 128$ and $L_c = 256$ are randomly generated and the resulting experimental $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$ is displayed

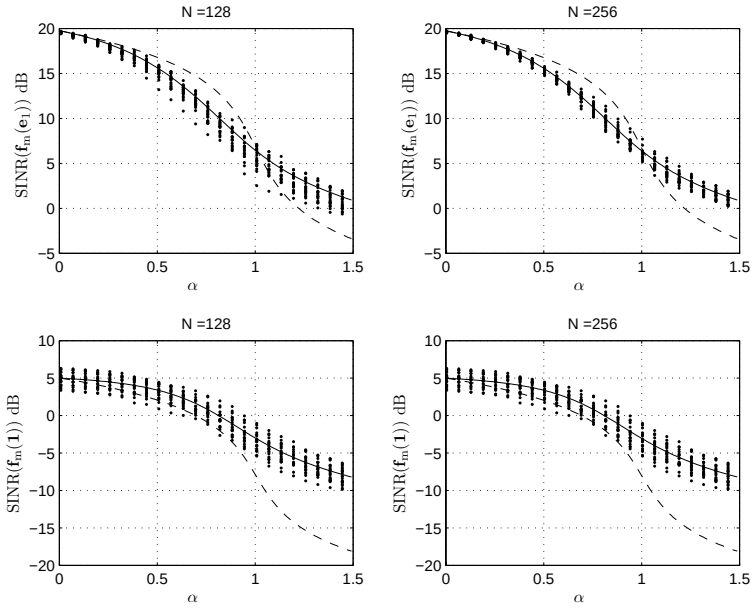

Figure 6.1: $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{1}))$ versus α in the downlink case.

versus the system load α for all sets of spreading codes, various choices of $\bar{\mathbf{h}}$, and both for the uplink and downlink cases. In the uplink case, the channel vectors $\mathbf{h}_k$ of length $L = 5$ are randomly and independently generated from a zero-mean circular multivariate Gaussian distribution $\mathcal{N}_C(\mathbf{0}, \frac{\sigma_{\mathbf{h}}^2}{L}\mathbf{I})$ with $\sigma_{\mathbf{h}}^2/\sigma^2 = 25$ dB. The so-obtained $\mathbf{h}_1$ is also used as the user-of-interest channel vector in the downlink case. The experimental results are compared with the asymptotic SINRs of Theorems 1 and 2. These asymptotic SINR curves are displayed using the dashed and solid lines, respectively.

Fig. 6.1 shows the results for the downlink case with the upper and lower subplots displaying $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{1}))$, respectively. As can be seen from Fig. 6.1, the experimental SINRs and analytical asymptotic SINRs of (6.95) are very close to each other, and as L_c increases, the spread of the experimental SINR values around corresponding analytical curves tends to decrease. The analytical SINR curves obtained from (6.50) show, however, that the results of Theorem 1 do not correctly predict the experimental SINR values in the downlink case. This observation validates our discussion in Section 6.4.

Fig. 6.2 shows the uplink counterpart of the results of Fig. 6.1. From Fig. 6.2, similar

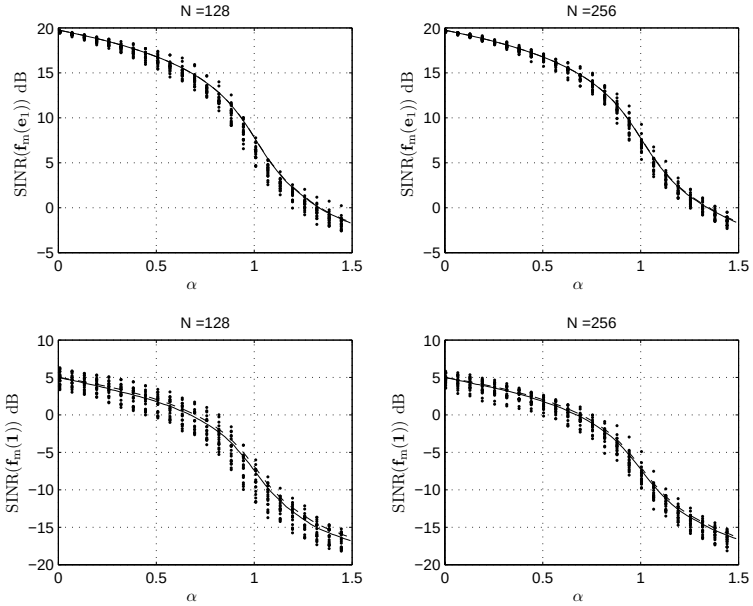

Figure 6.2: SINR($\mathbf{f}_{\rm m}(\mathbf{e}_1)$) and SINR($\mathbf{f}_{\rm m}(\mathbf{1})$) versus α in the uplink case.

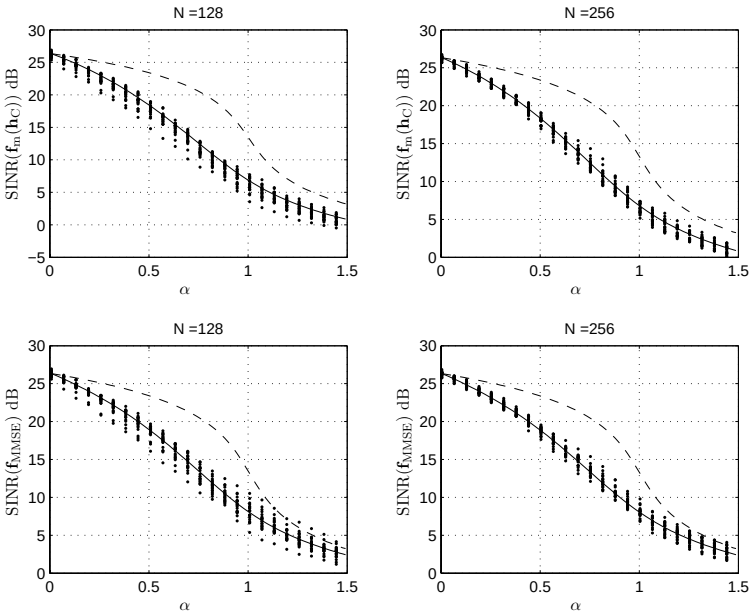

Figure 6.3: SINR($\mathbf{f}_{\rm m}(\mathbf{h}_{\rm C})$) and SINR($\mathbf{f}_{\rm MMSE}$) versus α in the downlink case.

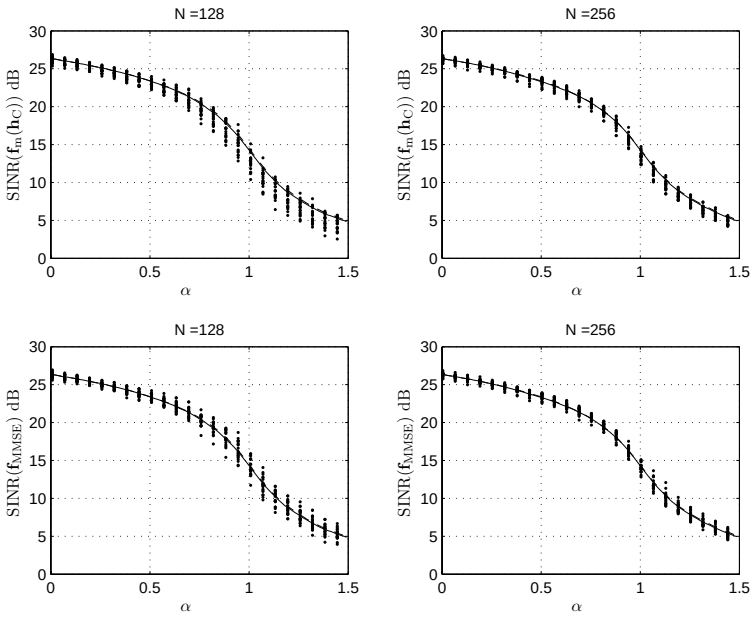

Figure 6.4: SINR($\mathbf{f}_{\rm m}(\mathbf{h}_{\rm C})$) and SINR($\mathbf{f}_{\rm MMSE}$) versus α in the uplink case.

observations can be made as in Fig. 6.1, except the fact that the asymptotic SINR curves obtained from (6.50) are very close to those obtained from (6.95). The latter fact observation the results of Theorem 4 stating that the asymptotic SINR expressions obtained based on the assumption of statistical independence of the columns of $\mathbf{C}_k$ are reliable approximations of the actual SINR values in the uplink case.

The downlink performance of the Capon and MMSE receivers are illustrated in Fig. 6.3, where the upper subplots display the experimental SINR($\mathbf{f}_{\rm m}(\mathbf{h}_{\rm C})$) while, for the sake of comparison, the experimental values of SINR($\mathbf{f}_{\rm MMSE}$) are shown in the lower subplots. The asymptotic SINR curves are also depicted in Fig. 6.3. Note that, due to (6.64), the dashed curve is in fact the asymptotic SINR($\mathbf{f}_{\rm MMSE}$) obtained from (6.49), while the solid curve is obtained from (6.95) wherein, due to (6.16) and (6.65), $\mathbf{h}_{\rm C} = \Omega\left(\mathbf{T}^{-1} + \mathbf{h}_1\mathbf{h}_1^H\right)$ is chosen. In the lower subplots, the asymptotic dashed and solid curves are obtained from (6.49) and (6.94), respectively. Similar to the observations made in Fig. 6.1, the asymptotic SINR results (6.94) and (6.95) of Theorem 2 predict the corresponding experimental SINRs with a very high accuracy which increases when L_c grows from 128 to 256. At the same time, the

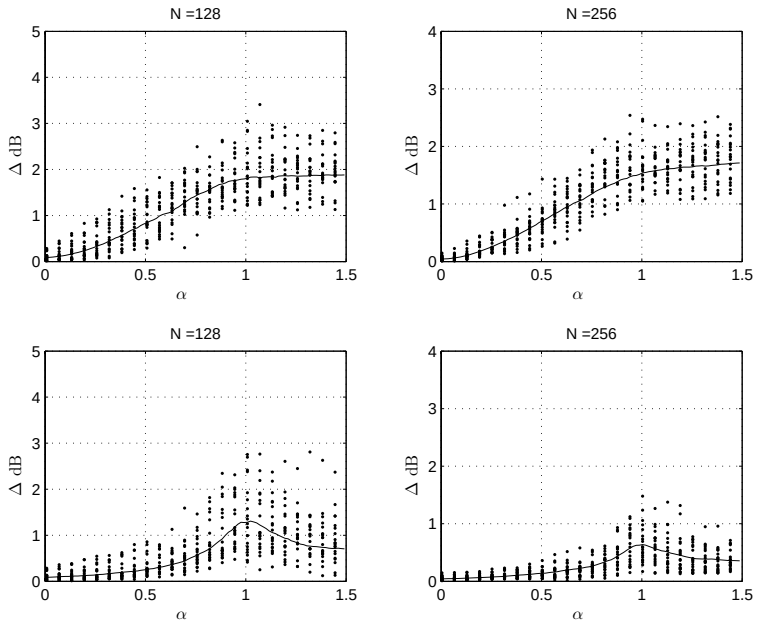

Figure 6.5: Δ versus α in the downlink (upper subplots) and uplink (lower subplots) cases.

asymptotic results (6.49) and (6.50) of Theorem 1 do not provide accurate enough estimates of the experimental SINRs in the downlink case.

Fig. 6.4 shows the uplink counterpart of the results of Fig. 6.3. It can be observed from this figure that the asymptotic curves of Theorem 1 are almost indistinguishable from the corresponding curves of Theorem 2. This validates the results of Theorem 4.

Fig. 6.5 shows $\Delta = \mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) - \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus α. The upper and lower subplots correspond to the downlink and uplink cases, respectively. The empirical average values of Δ taken over different realizations of the spreading codes are also shown by solid line. As can be observed from Fig. 6.5, the average Δ both in the uplink and downlink cases is less then 2 dB. However, this value is considerably smaller in the uplink case as compared to the downlink case. For instance, the average Δ barely reaches 0.5 dB in the uplink case with $L_c = 256$. This observation confirms the result in Section 6.4 stating that, under certain mild conditions, $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ approaches $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ in the uplink case.

In Figs. 6.6-6.8, the randomly generated spreading code of each user has the length of $L_c = 128$ and remains fixed over all simulation runs. Fig. 6.6 shows MSE between the

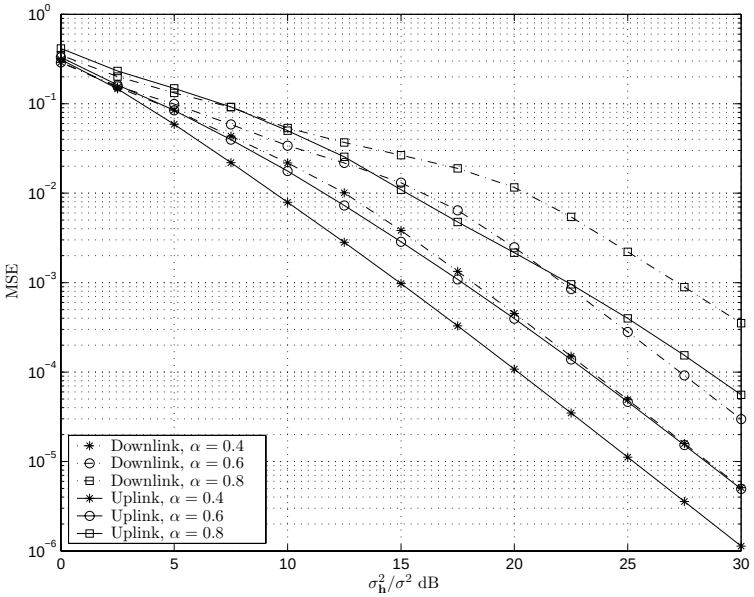

Figure 6.6: MSE between $\mathbf{h}_{\mathrm{C}}$ and the normalized version of $\mathbf{h}_1$ versus $\sigma^2_{\mathbf{h}}/\sigma^2$.

estimated channel vector $\mathbf{h}_{\mathrm{C}}$ and the normalized version of $\mathbf{h}_1$ versus $\sigma^2_{\mathbf{h}}/\sigma^2$ for both the uplink and downlink cases and three different values of α. The MSE curves are the results of averaging over 200 random realizations of the channel vectors $\mathbf{h}_k$ of length $L = 5$ that are independently generated from $\mathcal{N}_C(\mathbf{0}, \frac{\sigma^2_{\mathbf{h}}}{L}\mathbf{I})$. Note that the so-obtained realizations of $\mathbf{h}_1$ are used as the user-of-interest channel vectors in the downlink case. As can be observed from this figure, for each value of α the estimation performance in the uplink case is significantly better than in the downlink case.

Fig. 6.7 displays the experimental curves $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus $\sigma^2_{\mathbf{h}}/\sigma^2$ both in the uplink and downlink cases. It can be observed from this figure that, as the system load α increases, the gap between the uplink and downlink values of $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ grows.

Fig. 6.8 shows Δ versus the actual channel length $\tilde{L}$ for $L_c = 128$ and different values of α in the downlink case when the actual channel length is unknown to the blind Capon receiver $\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})$ while the channel vector is known to the MMSE receiver $\mathbf{f}_{\mathrm{MMSE}}$. In this figure, for any given $1 \leq \tilde{L} \leq 15$, the channel length $L = 15$ is used in (6.14) to obtain $\mathbf{h}_{\mathrm{C}}$, and, then, the so-obtained $\mathbf{h}_{\mathrm{C}}$ is exploited in (6.19) to compute $\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})$. Each simulated point in this

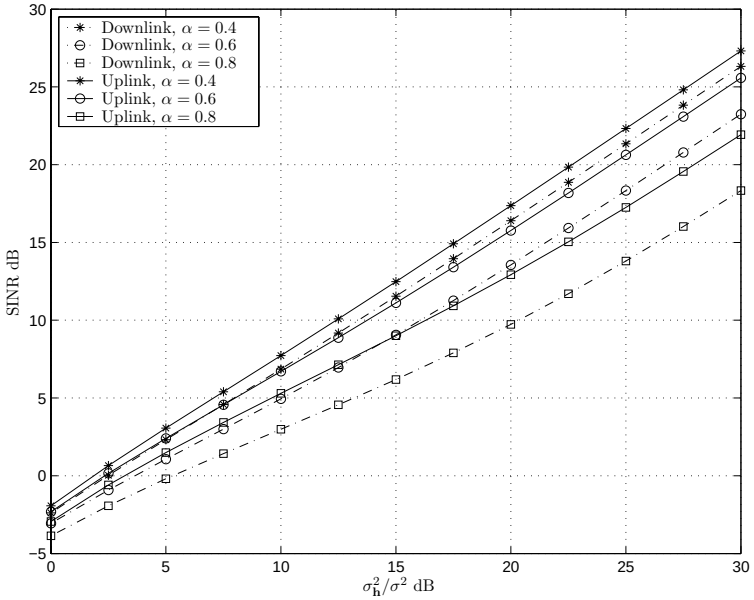

Figure 6.7: $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus $\sigma^2_{\mathbf{h}}/\sigma^2$.

figure is obtained as the average over 200 random realizations of the channel vector drawn from $\mathcal{N}_C(\mathbf{0}, \frac{\sigma^2_{\mathbf{h}}}{\tilde{L}}\mathbf{I})$ where $\sigma^2_{\mathbf{h}}/\sigma^2 = 25$ dB. As can be observed from Fig. 6.8, the smallest Δ corresponds to $\tilde{L} = 1$, even though this value of $\tilde{L}$ makes the largest difference between the actual and the presumed channel lengths. The latter observation further validates Corollary 1 stating that, if the actual channels are flat fading, then, regardless of the presumed channel length, the difference between $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ and $\mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}})$ converges to zero.

6.6 Chapter summary

The performance of blind MOE receivers in large CDMA systems has been studied in this chapter. In the first part of our analysis, we have used a simplifying assumption that the delayed versions of the spreading codes are statistically independent from each other to approximate the asymptotic SINR of the MOE receiver. Then, dropping the latter assumption, the true asymptotic SINR expressions have been derived and important special cases where the earlier approximate results closely follow the true results have been

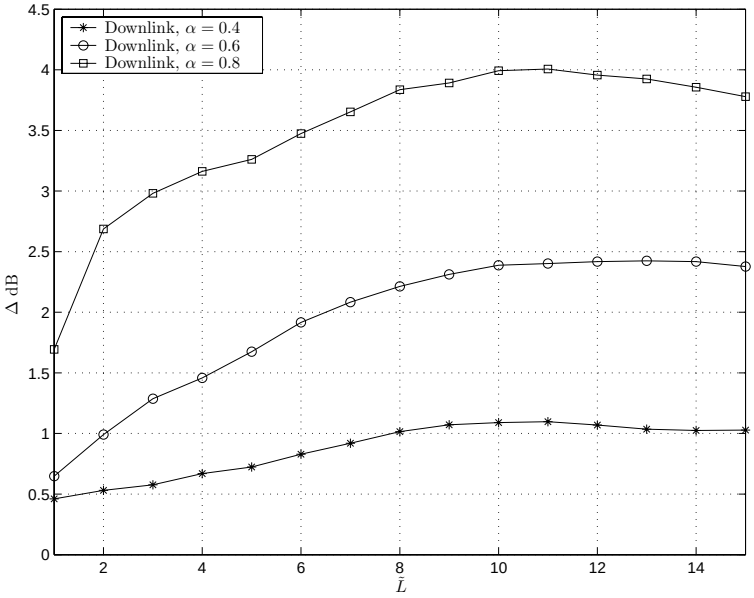

Figure 6.8: Δ versus the actual channel length $\tilde{L}$ when $L = 15$ is assumed in the Capon receiver.

identified. It has been shown that the performance of the optimized version of the MOE receiver (also referred to as the Capon receiver) is very close to that of the optimum MMSE receiver in the uplink mode. However, in the downlink mode, this property is no longer true.

Chapter 7

High SNR extension to asymptotic performance analysis of MOE receivers

7.1 Performance analysis

In this chapter we continue our studies in Chapter 6 by analyzing the performance of the MOE receiver in the high SNR regime. We need the following lemma to establish our main results.

Lemma 1: Let the $L \times L$ Hermitian Toeplitz matrix $\boldsymbol{\Upsilon}$ be defined as

$$[\boldsymbol{\Upsilon}]_{st} = \frac{1}{L_c} \sum_{n=1}^{L_c} \frac{e^{-j\frac{2\pi}{L_c}(t-s)(n-1)}}{\eta_n} \tag{7.1}$$

where $0 < d \leq \eta_n \leq D$ for $n = 1, \ldots, L_c$, and d and D are some constants. Then, $\boldsymbol{\Upsilon}$ is a positive definite matrix with

$$\lambda_{\min}(\boldsymbol{\Upsilon}) \geq D^{-1} \tag{7.2}$$

$$\lambda_{\max}(\boldsymbol{\Upsilon}) \leq d^{-1}. \tag{7.3}$$

Moreover, all entries of $\boldsymbol{\Upsilon}^{-1}$ have bounded norm.

Proof: Let $\boldsymbol{\Upsilon}_n$ be the $L \times L$ matrix with the following entries

$$[\boldsymbol{\Upsilon}_n]_{st} = \frac{1}{L_c} \frac{e^{-j\frac{2\pi}{L_c}(t-s)(n-1)}}{\eta_n}. \tag{7.4}$$

It can be readily shown that

$$\boldsymbol{\Upsilon}_n = \frac{1}{\eta_n} \mathbf{a}_n \mathbf{a}_n^H \tag{7.5}$$

where

$$\mathbf{a}_n = \frac{1}{\sqrt{L_c}} [1 \; e^{j\frac{2\pi}{L_c}(n-1)} \; e^{j\frac{4\pi}{L_c}(n-1)} \; \cdots \; e^{j\frac{2(L-1)\pi}{L_c}(n-1)}]^T. \tag{7.6}$$

From (7.1), (7.4), and (7.5) it follows that

$$\boldsymbol{\Upsilon} = \sum_{n=1}^{L_c} \boldsymbol{\Upsilon}_n = \sum_{n=1}^{L_c} \frac{1}{\eta_n} \mathbf{a}_n \mathbf{a}_n^H. \tag{7.7}$$

Let

$$\boldsymbol{\Upsilon}_{\min} = \frac{1}{D} \sum_{n=1}^{L_c} \mathbf{a}_n \mathbf{a}_n^H \tag{7.8}$$

$$\boldsymbol{\Upsilon}_{\max} = \frac{1}{d} \sum_{n=1}^{L_c} \mathbf{a}_n \mathbf{a}_n^H. \tag{7.9}$$

Note that $\sum_{n=1}^{L_c} \mathbf{a}_n \mathbf{a}_n^H = \mathbf{A}\mathbf{A}^H$ where

$$\mathbf{A} = [\mathbf{a}_1 \; \mathbf{a}_2 \; \cdots, \mathbf{a}_{L_c}]. \tag{7.10}$$

From (7.6) and (7.10), we have

$$[\mathbf{A}\mathbf{A}^H]_{st} = \frac{1}{L_c} \sum_{n=1}^{L_c} e^{j\frac{2\pi}{L_c}(s-t)(n-1)} = \delta_{st} \tag{7.11}$$

or, equivalently, $\mathbf{A}\mathbf{A}^H = \mathbf{I}$. From the latter fact, it immediately follows that

$$\boldsymbol{\Upsilon}_{\min} = \frac{1}{D} \mathbf{I} \tag{7.12}$$

$$\boldsymbol{\Upsilon}_{\max} = \frac{1}{d} \mathbf{I}. \tag{7.13}$$

For any arbitrary $L \times 1$ vector $\mathbf{u}$ we have

$$\begin{aligned}\mathbf{u}^H \left(\boldsymbol{\Upsilon} - \frac{1}{D}\mathbf{I} \right) \mathbf{u} &= \mathbf{u}^H \left(\sum_{n=1}^{L_c} \left(\frac{1}{\eta_n} - \frac{1}{D} \right) \mathbf{a}_n \mathbf{a}_n^H \right) \mathbf{u} \\ &= \sum_{n=1}^{L_c} \left(\frac{1}{\eta_n} - \frac{1}{D} \right) |\mathbf{u}^H \mathbf{a}_n|^2 \geq 0 \end{aligned} \tag{7.14}$$

where the first line of (7.14) follows from (7.7), (7.8) and (7.12). Inequality (7.14) implies that $\boldsymbol{\Upsilon} - \frac{1}{D}\mathbf{I}$ is a positive semi-definite matrix. Therefore, $\boldsymbol{\Upsilon}$ is a positive definite matrix with $\lambda_{\min}(\boldsymbol{\Upsilon}) \geq D^{-1}$. Using similar steps as in (7.14), it can be shown that $\frac{1}{d}\mathbf{I} - \boldsymbol{\Upsilon}$ is also a positive semi-definite matrix, and, hence, $\lambda_{\max}(\boldsymbol{\Upsilon}) \leq d^{-1}$. Finally, as all entries of $\boldsymbol{\Upsilon}$ are norm-bounded, all the minors of $\boldsymbol{\Upsilon}$ have bounded norms. Using this fact along with the fact that the determinant of $\boldsymbol{\Upsilon}$ is nonzero, it follows that all entries of $\boldsymbol{\Upsilon}^{-1}$ are norm-bounded.

□

Before proceeding with our analysis, let us assume that

$$|h_k\big(e^{j\frac{2\pi}{L_c}(n-1)}\big)|^2 \geq b \qquad k = 1, \ldots, K \qquad n = 1, \ldots, L_c \tag{7.15}$$

where $b > 0$ is a constant. Note that from (6.67) it follows that the sporadic (k, n)s for which $|h_k\big(e^{j\frac{2\pi}{L_c}(n-1)}\big)|^2 = 0$ do not have any impact on the value of ϕ_l and, hence, they can be discarded from (6.67).

From (6.16) and (6.21), it can be observed that the high SNR performances of the MOE receiver and the Capon channel estimation technique are essentially dictated by the properties of $\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1$ in the high SNR regime. These properties are derived in the following theorem.

Theorem 1: For $\alpha < 1$, we have

$$\lim_{\sigma^2 \to 0} \left(\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 \right)^{-1} \overset{e.i.p.}{\longrightarrow} \mathbf{0} \tag{7.16}$$

$$\lim_{\sigma^2 \to 0} \sigma^2 \underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \check{\boldsymbol{\Psi}} \overset{e.i.p.}{\longrightarrow} \mathbf{0} \tag{7.17}$$

where $\check{\boldsymbol{\Psi}}$ is a positive definite Hermitian Toeplitz matrix with

$$[\check{\boldsymbol{\Psi}}]_{1m} = \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{e^{-j\frac{2\pi}{L_c}(m-1)(l-1)}}{\check{\psi}_l} \qquad m = 1, \ldots, L \tag{7.18}$$

and $\check{\psi}_l$ is a unique real and positive solution to

$$\check{\psi}_l = 1 + \frac{1}{L_c}\sum_{k=2}^{K} \frac{\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2}{\frac{1}{L_c}\sum_{n=1}^{L_c}\frac{\left|h_k(e^{j\frac{2\pi}{L_c}(n-1)})\right|^2}{\check{\psi}_n}}, \qquad l = 1, \ldots, L_c. \tag{7.19}$$

Moreover,

$$\max\left\{1, \frac{b}{B(1-\alpha)}\right\} \leq \check{\psi}_l \leq \frac{B}{b(1-\alpha)}, \qquad l = 1, \ldots, L_c. \tag{7.20}$$

For $\alpha > 1$, we have

$$\lim_{\sigma^2 \to 0} \underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \check{\boldsymbol{\Phi}} \xrightarrow{e.i.p.} \mathbf{0} \tag{7.21}$$

where $\check{\boldsymbol{\Phi}}$ is a positive definite Hermitian Toeplitz matrix with

$$[\check{\boldsymbol{\Phi}}]_{1m} = \frac{1}{L_c}\sum_{l=1}^{L_c} \frac{e^{-j\frac{2\pi}{L_c}(m-1)(l-1)}}{\check{\phi}_l}, \qquad m = 1, \ldots, L \tag{7.22}$$

and $\check{\phi}_l$ is the unique real and positive solution to

$$\check{\phi}_l = \frac{1}{L_c}\sum_{k=2}^{K} \frac{\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2}{1 + \frac{1}{L_c}\sum_{n=1}^{L_c}\frac{\left|h_k(e^{j\frac{2\pi}{L_c}(n-1)})\right|^2}{\check{\phi}_n}}, \qquad l = 1, \ldots, L_c. \tag{7.23}$$

Moreover,

$$\frac{b^2(\alpha-1)}{B} \leq \check{\phi}_l \leq \frac{B^2(\alpha-1)}{b}. \tag{7.24}$$

Proof: Before presenting the proof of Theorem 1, some remarks are in order. First, using (7.15) in (6.67) we have

$$\sigma^2 + \kappa b \leq \phi_l \leq \sigma^2 + \kappa B, \qquad l = 1, \ldots, L_c \tag{7.25}$$

where

$$\kappa = \frac{1}{L_c}\sum_{k=2}^{K} \frac{1}{1 + \frac{1}{L_c}\sum_{n=1}^{L_c}\frac{\left|h_k(e^{j\frac{2\pi}{L_c}(n-1)})\right|^2}{\phi_n}} \leq \alpha. \tag{7.26}$$

Therefore, for any $l_1, l_2 \in \{1, \ldots, L_c\}$, we have

$$\frac{b}{B} \leq \frac{\sigma^2 + \kappa b}{\sigma^2 + \kappa B} \leq \frac{\phi_{l_1}}{\phi_{l_2}} \leq \frac{\sigma^2 + \kappa B}{\sigma^2 + \kappa b} \leq \frac{B}{b}. \tag{7.27}$$

Let $\psi_l = \dfrac{\phi_l}{\sigma^2}$. As ϕ_l is a differentiable function of σ^2, both $\check{\phi}_l = \lim_{\sigma^2 \to 0} \phi_l$ and $\check{\psi}_l = \lim_{\sigma^2 \to 0} \psi_l$ exist. Note that $\check{\phi}_l$, ψ_l, and $\check{\psi}_l$ are real and nonnegative and $\check{\psi}_l$ may be equal to $+\infty$. From (7.27), we have

$$\frac{b}{B} \leq \frac{\check{\phi}_{l_1}}{\check{\phi}_{l_2}} = \frac{\psi_{l_1}}{\psi_{l_2}} \leq \frac{B}{b}, \qquad l_1, l_2 \in \{1, \ldots, L_c\}. \tag{7.28}$$

Moreover, it follows from (7.27) that if there is a bounded $\check{\psi}_{l^*}$, then all $\check{\psi}_l$ are bounded and we have

$$\frac{b}{B} \leq \frac{\check{\psi}_{l_1}}{\check{\psi}_{l_2}} \leq \frac{B}{b}, \qquad l_1, l_2 \in \{1, \ldots, L_c\}. \tag{7.29}$$

In turn, if there is a $\check{\psi}_{l^*} = +\infty$, then $\check{\psi}_l = +\infty$ for $l = 1, \ldots, L_c$.

From (6.26) and (6.65)-(6.67), it directly follows that as $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha$

$$\sigma^2 \underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \boldsymbol{\Psi} \xrightarrow{e.i.p.} \mathbf{0} \tag{7.30}$$

where $\boldsymbol{\Psi} = \sigma^2 \mathbf{T}$ is the positive definite Hermitian Toeplitz matrix with

$$[\boldsymbol{\Psi}]_{1m} = \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{e^{-j\frac{2\pi}{L_c}(m-1)(l-1)}}{\psi_l} \qquad m = 1, \ldots, L \tag{7.31}$$

and ψ_l is the unique real and positive solution to

$$\psi_l = 1 + \frac{1}{L_c} \sum_{k=2}^{K} \frac{\left| h_k(e^{j\frac{2\pi}{L_c}(l-1)}) \right|^2}{\omega_k} \qquad l = 1, \ldots, L_c \tag{7.32}$$

where

$$\omega_k = \sigma^2 + \frac{1}{L_c} \sum_{n=1}^{L_c} \frac{\left| h_k(e^{j\frac{2\pi}{L_c}(n-1)}) \right|^2}{\psi_n}. \tag{7.33}$$

From (7.32), it follows that $\psi_l \geq 1$. Dividing both sides of (7.32) by ψ_l, summing the resulting equations over l, and dividing both sides by L_c, we have

$$\begin{aligned} 1 &= \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\psi_l} + \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{L_c} \sum_{k=2}^{K} \frac{\left| h_k(e^{j\frac{2\pi}{L_c}(l-1)}) \right|^2}{\psi_l \omega_k} \\ &= \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\psi_l} + \frac{1}{L_c} \sum_{k=2}^{K} \frac{1}{\omega_k} \cdot \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{\left| h_k(e^{j\frac{2\pi}{L_c}(l-1)}) \right|^2}{\psi_l} \\ &= \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\psi_l} + \frac{1}{L_c} \sum_{k=2}^{K} \frac{\omega_k - \sigma^2}{\omega_k}. \end{aligned} \tag{7.34}$$

Consider (7.34) in the case $\sigma^2 \to 0$. Regardless of the value of $\check{\omega}_k = \lim_{\sigma^2 \to 0} \omega_k$, we have that

$$\lim_{\sigma^2 \to 0} \frac{1}{L_c} \sum_{k=2}^{K} \frac{\omega_k - \sigma^2}{\omega_k} \leq \alpha. \tag{7.35}$$

Therefore, it follows from (7.34) and (7.35) that

$$\frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\check{\psi}_l} \geq 1 - \alpha. \tag{7.36}$$

Hence, if $\alpha < 1$, then $\check{\psi}_l,\ l = 1, \ldots, L_c$ are bounded scalars, and, consequently,

$$\check{\omega}_k = \frac{1}{L_c} \sum_{n=1}^{L_c} \frac{\left| h_k(e^{j\frac{2\pi}{L_c}(n-1)}) \right|^2}{\check{\psi}_n}. \tag{7.37}$$

Note that in such a case, $\check{\omega}_k > 0$. Considering (7.32) in the case $\sigma^2 \to 0$ and using (7.37), it follows that $\check{\psi}_l, l = 1, \ldots, L_c$ are the unique positive solutions to (7.19). In turn, (7.17) and (7.18) follow by taking the limit of (7.30) and (7.31) for $\sigma^2 \to 0$.

To prove (7.20), note that according to (7.29) for any arbitrary choice of $l^* \in \{1, \ldots, L_c\}$,

$$\frac{B}{b} \cdot \frac{1}{\check{\psi}_{l^*}} \geq \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\check{\psi}_l}. \tag{7.38}$$

The right-hand side inequality in (7.20) is directly obtained from (7.36) and (7.38). The left-hand side inequality in (7.20) can be proved as follows. Since $\check{\omega}_k > 0$, we have

$$\lim_{\sigma^2 \to 0} \frac{1}{L_c} \sum_{k=2}^{K} \frac{\omega_k}{\sigma^2 + \omega_k} = \alpha. \tag{7.39}$$

Hence, for $\alpha < 1$ (7.36) can be replaced by

$$\frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\check{\psi}_l} = 1 - \alpha. \tag{7.40}$$

Again, (7.29) can be used to obtain

$$\frac{b}{B} \cdot \frac{1}{\check{\psi}_{l^*}} \leq \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\check{\psi}_l} \tag{7.41}$$

for $l^* \in \{1, \ldots, L_c\}$. Using (7.40) and (7.41) along with the fact that $\check{\psi}_l \geq 1$, the left-hand side inequality of (7.20) follows. Note that according to (7.20) and Lemma 1, $\check{\mathbf{\Psi}}$ is a positive

definite matrix, and, hence, is invertible. Now, we are able to prove (7.16) as follows. We have

$$\left(\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1\right)^{-1} = \sigma^2 \left(\sigma^2 \underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1\right)^{-1}. \tag{7.42}$$

Therefore, from (7.17) we have

$$\lim_{\sigma^2 \to 0} \left(\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1\right)^{-1} - \lim_{\sigma^2 \to 0} \sigma^2 \check{\boldsymbol{\Psi}}^{-1} \stackrel{e.i.p.}{\longrightarrow} \mathbf{0}. \tag{7.43}$$

As, according to Lemma 1 the entries of $\check{\boldsymbol{\Psi}}^{-1}$ are norm-bounded, (7.16) is immediately obtained from (7.43).

Before proving (7.21)-(7.24), we need to show that for $\alpha > 1$

$$\check{\psi}_l = +\infty, \qquad l = 1, \ldots, L_c. \tag{7.44}$$

Equation (7.44) can be proved by contradiction as follows. Assume that there exists $l^* \in \{1, \ldots, L_c\}$ such that $\check{\psi}_{l^*}$ is bounded for $\alpha > 1$. In such a case, from (7.29) we have that all $\check{\psi}_l,\ l = 1, \ldots, L_c$ are bounded, and, therefore, $\check{\omega}_k > 0$ for $k = 1, \ldots, K$. This, in turn, results in (7.39) and (7.40). In particular, from the latter equation and the fact that $\alpha > 1$, it follows that

$$\frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\check{\psi}_l} < 0. \tag{7.45}$$

Inequality (7.45) is clearly wrong due to the fact that $\check{\psi}_l > 1$ for $l = 1, \ldots, L_c$. The latter fact immediately proves (7.44).

We can use (7.44) to prove (7.24) for $\alpha > 1$. Dividing both sides of (6.67) by ϕ_l, summing up the resulting equations over l, and, dividing both sides of the so-obtained equation by L_c, we have

$$1 = \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{1}{\psi_l} + \frac{1}{L_c} \sum_{k=2}^{K} \frac{\gamma_k}{1 + \gamma_k} \tag{7.46}$$

where

$$\gamma_k = \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2}{\phi_l}. \tag{7.47}$$

Assuming that $\sigma^2 \to 0$ and using (7.44) in (7.46), we get

$$1 = \frac{1}{L_c} \sum_{k=2}^{K} \frac{\check{\gamma}_k}{1 + \check{\gamma}_k} \tag{7.48}$$

where

$$\check{\gamma}_k = \frac{1}{L_c} \sum_{l=1}^{L_c} \frac{\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2}{\check{\phi}_l}. \tag{7.49}$$

Note that $\check{\phi}_l > 0$ for $l = 1, \ldots, L_c$: If there is $\check{\phi}_{l^*} = 0$, then, due to (7.28), $\check{\phi}_l = 0$ for all $l = 1, \ldots, L_c$. It immediately follows from the latter fact and (7.47) that $\check{\gamma}_k = +\infty$ for $k = 2, \ldots, K$, and, consequently, the right-hand side of (7.48) converges to α. This contradicts to our assumption that $\alpha > 1$. Now, consider an arbitrary choice of $l^* \in \{1, \ldots, L_c\}$. According to (7.28), we have

$$\frac{b}{B\check{\phi}_{l^*}} \leq \frac{1}{\check{\phi}_l} \leq \frac{B}{b\check{\phi}_{l^*}}, \qquad l = 1, \ldots, L_c. \tag{7.50}$$

Using (7.49) and (7.50) along with the fact that $b \leq \left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|^2 \leq B$, it follows that

$$\frac{b^2}{B\check{\phi}_{l^*}} \leq \check{\gamma}_k \leq \frac{B^2}{b\check{\phi}_{l^*}}. \tag{7.51}$$

As $\frac{\check{\gamma}_k}{1+\check{\gamma}_k}$ is a strictly increasing function of $\check{\gamma}_k$, from (7.51) we obtain

$$\frac{b^2}{b^2 + B\check{\phi}_{l^*}} \leq \frac{\check{\gamma}_k}{1+\check{\gamma}_k} \leq \frac{B^2}{B^2 + b\check{\phi}_{l^*}}. \tag{7.52}$$

Using (7.52) in (7.48) yields

$$\frac{\alpha b^2}{b^2 + B\check{\phi}_{l^*}} \leq 1 \leq \frac{\alpha B^2}{B^2 + b\check{\phi}_{l^*}}. \tag{7.53}$$

Inequalities (7.24) can be directly obtained from (7.53).

To prove (7.23), just note that (7.24) implies that $\check{\phi}_l > 0$ for $l = 1, \ldots, L_c$. Therefore, the effect of σ^2 in the right-hand side of (6.67) becomes negligible as σ^2 converges to zero and, consequently, $\check{\phi}_l$ follow the equation (7.23) for $l = 1, \ldots, L_c$. Finally, denoting $\check{\mathbf{\Phi}} = \lim_{\sigma^2 \to 0} \mathbf{T}$, (7.21) and (7.22) directly follow from (6.65) and (6.66), respectively. □

Note that from (7.20) and (7.24) along with Lemma 1, we have

$$\lambda_{\min}(\check{\mathbf{\Psi}}) \geq \frac{b(1-\alpha)}{B} \tag{7.54}$$

$$\lambda_{\max}(\check{\mathbf{\Psi}}) \leq \min\left\{1, \frac{B(1-\alpha)}{b}\right\} \tag{7.55}$$

$$\lambda_{\min}(\check{\mathbf{\Phi}}) \geq \frac{b}{B^2(\alpha-1)} \tag{7.56}$$

$$\lambda_{\max}(\check{\mathbf{\Phi}}) \leq \frac{B}{b^2(\alpha-1)}. \tag{7.57}$$

The following results can be directly obtained from Theorem 1.

Corollary 1: Assume that $\alpha < 1$. Then, as $\sigma^2 \to 0$, there exists $\theta \in [0, 2\pi)$ such that

$$\mathbf{h}_{\mathrm{C}} \xrightarrow{e.i.p.} e^{j\theta}\frac{\mathbf{h}_1}{\|\mathbf{h}_1\|}. \tag{7.58}$$

Moreover, for any arbitrary $\bar{\mathbf{h}}$ such that $\bar{\mathbf{h}}^H\mathbf{h}_1 \neq 0$, $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$ goes to infinity.

Note that Equation (7.58) directly follows from using (7.16) in (6.14). Equation (7.16) can also be used in (6.21) to verify the fact that $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \to \infty$. The above properties of the MOE receiver (6.19) and the Capon channel estimate (6.14) can be further elaborated as follows.

- It has been shown in [78] that if

$$[\underline{\boldsymbol{\mathcal{C}}}_1 \quad \mathbf{W}_1] \text{ is a full column-rank matrix} \tag{7.59}$$

where $\mathbf{W}_1 = [\mathbf{w}_2, \ldots, \mathbf{w}_K]$, then, for $\sigma^2 \to 0$, $\mathbf{h}_{\mathrm{C}}$ converges to a scaled version of $\mathbf{h}_1$. Corollary 2 proves that for $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha$, the identifiability condition (7.59) can be replaced by the much simpler condition $\alpha < 1$. That is certainly a desirable property since the verification of (7.59) can be prohibitively difficult when L_c and K are large. Note that such a simplification of the identifiability condition is also intuitively justifiable: Since the spreading sequences of users are independently drawn, as long as $[\underline{\boldsymbol{\mathcal{C}}}_1 \quad \mathbf{W}_1]$ is a tall matrix (i.e., $\alpha < 1$), it is likely that $[\underline{\boldsymbol{\mathcal{C}}}_1 \quad \mathbf{W}_1]$ is full-column rank. However, Corollary 1 is a rigorous proof for the above conjecture showing that if $\alpha < 1$, then, for the asymptotic scenario of our concern, $\mathbf{h}_1$ is uniquely identifiable.

- As $\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = \infty$ for $\alpha < 1$, it follows that in the absence of noise the general blind MOE receiver (6.19) is able to completely suppress the effect of MAI regardless of the interferer powers. It is an interesting result which shows that even without knowing the signature of the user-of-interest $\mathbf{w}_1$ or having an estimate of the channel vector $\mathbf{h}_1$, the receiver (6.19) still has the near-far resistance property. This property is due to the fact that $\mathbf{w}_1$ is equal to $\underline{\boldsymbol{\mathcal{C}}}_1 \mathbf{h}_1$ where $\underline{\boldsymbol{\mathcal{C}}}_1$ is a known matrix. The MOE receiver (6.19) effectively uses the known structure of $\mathbf{w}_1$ to suppress the interfering signals having the signatures $\mathbf{w}_k = \underline{\boldsymbol{\mathcal{C}}}_k \mathbf{h}_k$, for $k = 2, \ldots, K$.

Note also that if the actual channel order is $\tilde{L} < L$, i.e., $\mathbf{h}_1 = [\tilde{\mathbf{h}}_1^T \quad \mathbf{0}^T]^T$ where $\tilde{\mathbf{h}}_1$ is a vector of length $\tilde{L}$, then, (7.16) still holds and it can be used in (6.14) to obtain (7.58). The latter observation shows that if $\alpha < 1$, then the Capon channel estimation technique is insensitive to the channel order overestimation in the high SNR regime.

It is known that if the number of active users exceeds the signature length, i.e., $\alpha > 1$, then, in the absence of noise, the MMSE receiver is not able to completely suppress the effect of MAI, and, hence, the resulting SINR does not go to infinity (see, e.g., [50], [79]). It is our intuitive conjecture that the similar property should also hold for the MOE receiver (6.19). The following corollary verifies such a conjecture.

Corollary 2: Assuming that $\alpha > 1$, we have

$$\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\bar{\mathbf{h}}^H \check{\boldsymbol{\Phi}}^{-1} \bar{\mathbf{h}}} \xrightarrow{e.i.p.} \mathbf{0} \tag{7.60}$$

and, therefore,

$$\frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\|\bar{\mathbf{h}}\|^2} \cdot \frac{b}{B^2(\alpha - 1)} \leq \lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \leq \frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\|\bar{\mathbf{h}}\|^2} \cdot \frac{B}{b^2(\alpha - 1)}. \tag{7.61}$$

If $b = B$, then

$$\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = \frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\|\bar{\mathbf{h}}\|^2} \cdot \frac{1}{B(\alpha - 1)} \tag{7.62}$$

and

$$\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})) = \frac{1}{\alpha - 1}. \tag{7.63}$$

Proof: Equation (7.60) straightforwardly follows from (6.21) and (7.21). To prove (7.61), note that from (7.56) and (7.57) we have

$$\frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\bar{\mathbf{h}}^H\check{\mathbf{\Phi}}^{-1}\bar{\mathbf{h}}} \leq \frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\|\bar{\mathbf{h}}^H\|^2\lambda_{\min}(\check{\mathbf{\Phi}}^{-1})} \leq \frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\|\bar{\mathbf{h}}\|^2} \cdot \frac{B}{b^2(\alpha-1)} \tag{7.64}$$

$$\frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\bar{\mathbf{h}}^H\check{\mathbf{\Phi}}^{-1}\bar{\mathbf{h}}} \geq \frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\|\bar{\mathbf{h}}^H\|^2\lambda_{\max}(\check{\mathbf{\Phi}}^{-1})} \geq \frac{|\bar{\mathbf{h}}^H\mathbf{h}_1|^2}{\|\bar{\mathbf{h}}\|^2} \cdot \frac{b}{B^2(\alpha-1)}. \tag{7.65}$$

Equation (7.60) along with inequalities (7.64) and (7.65) imply (7.61). If $b = B$, or, equivalently, $\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|$ is constant for all $k = 1, \ldots, K$ and $l = 1, \ldots, L_c$, then (7.62) follows from the equality of the upper and lower bounds in (7.61). Note that $\left|h_k(e^{j\frac{2\pi}{L_c}(l-1)})\right|$ does not change with k and l if the received user powers are equal and the channels are single-path, that is,

$$\mathbf{h}_k = h_{k,q_k}\mathbf{e}_{q_k}, \qquad k = 1, \ldots, K. \tag{7.66}$$

where $q_k \in \{0, 1, \ldots, L-1\}$. In such a case, we have $\|\mathbf{h}_k\|^2 = |h_{k,q_k}|^2 = B$, and, moreover, as shown in Chapter 6, $\mathbf{h}_{\mathrm{C}}$ converges to a scaled version of $\mathbf{h}_1 = h_{1,q_1}\mathbf{e}_{q_1}$. Using these results in (7.62), equation (7.63) follows. □

Note that, if $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha > 1$, and the received user powers are equal and the channels are single-path, then [79]

$$\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) = \frac{1}{\alpha - 1} \tag{7.67}$$

This, along with (7.63) is consistent with the results of Corollary 1 in Chapter 6 which indicates that if (6.96) holds true, then the $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}})) - \mathrm{SINR}(\mathbf{f}_{\mathrm{MMSE}}) \to 0$.

It should be mentioned that if $B \gg b$, then the derived bounds in (7.61) may become loose. However, for instance, in the downlink transmission scheme where the channel coefficient corresponding to the line-of-sight is much larger than the other coefficients and the transmitted power to the users are equal, it can be shown that B and b are close to each other, and, hence, (7.61) offers reasonably tight bounds on the asymptotic value of $\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$.

We can also use the results of Theorem 1 to analyze other performance measures of the receiver (6.19) such as the efficiency and the asymptotic efficiency. Note that the efficiency of the receiver vector $\mathbf{f}$, denoted here by $\mathrm{C}(\mathbf{f})$, is the ratio of the achieved SINR to the SINR when there is no interference [79]. Moreover, the asymptotic efficiency of the receiver vector $\mathbf{f}$ is defined as $\eta(\mathbf{f}) = \lim_{\sigma^2 \to 0} \mathrm{C}(\mathbf{f})$. Then, we have

Theorem 2: As $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha$, the efficiency of the blind MOE receiver (6.19) satisfies

$$\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \frac{\sigma^2 \|\bar{\mathbf{h}}\|^2}{\bar{\mathbf{h}}^H \mathbf{T}^{-1} \bar{\mathbf{h}}} \xrightarrow{i.p.} 0 \tag{7.68}$$

and is bounded by

$$\frac{\sigma^2}{\sigma^2 + \alpha B} \leq \mathrm{C}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \leq 1. \tag{7.69}$$

Moreover, for $\alpha < 1$ we have

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \frac{\|\bar{\mathbf{h}}\|^2}{\bar{\mathbf{h}}^H \check{\mathbf{\Psi}}^{-1} \bar{\mathbf{h}}} \xrightarrow{i.p.} 0 \tag{7.70}$$

and

$$\frac{b(1-\alpha)}{B} \leq \eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \leq \min\left\{1, \frac{B(1-\alpha)}{b}\right\}. \tag{7.71}$$

In turn, if $\alpha > 1$, then

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \xrightarrow{i.p.} 0. \tag{7.72}$$

Proof: From (6.21), we have

$$\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = \frac{\bar{\mathbf{h}}^H \left(\underline{\boldsymbol{\mathcal{C}}}_1^H (\underline{\tilde{\mathbf{w}}} \underline{\tilde{\mathbf{w}}}_1^H + \sigma^2 \mathbf{I})^{-1} \underline{\boldsymbol{\mathcal{C}}}_1\right)^{-1} \bar{\mathbf{h}}}{\bar{\mathbf{h}}^H (\underline{\boldsymbol{\mathcal{C}}}_1^H \mathbf{A}^{-1} \underline{\boldsymbol{\mathcal{C}}}_1)^{-1} \bar{\mathbf{h}}}. \tag{7.73}$$

Using a similar argument as in the proof of Proposition 1 in Chapter 6, it can be easily shown that as L_c and K go to infinity while L remains constant, we have

$$\underline{\boldsymbol{\mathcal{C}}}_1^H (\underline{\tilde{\mathbf{w}}}_1 \underline{\tilde{\mathbf{w}}}_1^H + \sigma^2 \mathbf{I})^{-1} \underline{\boldsymbol{\mathcal{C}}}_1 - \frac{1}{\sigma^2} \underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\boldsymbol{\mathcal{C}}}_1 \xrightarrow{e.a.s.} 0. \tag{7.74}$$

From [19, Lemma 1], we also have

$$\underline{\boldsymbol{\mathcal{C}}}_1^H \underline{\boldsymbol{\mathcal{C}}}_1 \xrightarrow{e.a.s.} \mathbf{I}. \tag{7.75}$$

Using (6.65), (7.74) and (7.75) in (7.73), Equation (7.68) follows. From (7.68), it follows that as $L_c \to \infty$ with $\frac{K}{L_c} \to \alpha$,

$$\frac{\sigma^2}{\lambda_{\max}(\mathbf{T}^{-1})} \leq \mathrm{C}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \leq \frac{\sigma^2}{\lambda_{\min}(\mathbf{T}^{-1})} \tag{7.76}$$

or, equivalently,

$$\sigma^2\lambda_{\min}(\mathbf{T}) \leq \mathrm{C}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \leq \sigma^2\lambda_{\max}(\mathbf{T}). \tag{7.77}$$

From Lemma 1 and (7.25)-(7.26), it follows that

$$\begin{aligned} \lambda_{\min}(\mathbf{T}) &\geq \frac{1}{\sigma^2 + \alpha B} \\ \lambda_{\max}(\mathbf{T}) &\leq \frac{1}{\sigma^2}. \end{aligned} \tag{7.78}$$

Using the latter results in (7.77), inequalities (7.69) can be obtained.

To prove (7.71), note that

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = \lim_{\sigma^2\to 0} \frac{\bar{\mathbf{h}}^H \left(\underline{\boldsymbol{\mathcal{C}}}_1^H(\tilde{\underline{\mathbf{w}}}_1\tilde{\underline{\mathbf{w}}}_1^H + \sigma^2\mathbf{I})^{-1}\underline{\boldsymbol{\mathcal{C}}}_1\right)^{-1}\bar{\mathbf{h}}}{\bar{\mathbf{h}}^H(\underline{\boldsymbol{\mathcal{C}}}_1^H\mathbf{A}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}}}. \tag{7.79}$$

Hence, it follows from (7.74) and (7.75) that

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \lim_{\sigma^2\to 0} \frac{\|\bar{\mathbf{h}}\|^2}{\bar{\mathbf{h}}^H(\sigma^2\underline{\boldsymbol{\mathcal{C}}}_1^H\mathbf{A}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}}} \xrightarrow{a.s.} 0. \tag{7.80}$$

Using (7.17) in (7.80), Equation (7.70) follows. We directly obtain from (7.70) that

$$\lambda_{\min}(\check{\boldsymbol{\Psi}}) \leq \eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \leq \lambda_{\max}(\check{\boldsymbol{\Psi}}). \tag{7.81}$$

Inequalities (7.71) follow by using (7.54) and (7.55) in (7.81). Finally, to prove (7.72) note that (7.80) can be rewritten as

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \lim_{\sigma^2\to 0} \frac{\sigma^2\|\bar{\mathbf{h}}\|^2}{\bar{\mathbf{h}}^H(\underline{\boldsymbol{\mathcal{C}}}_1^H\mathbf{A}^{-1}\underline{\boldsymbol{\mathcal{C}}}_1)^{-1}\bar{\mathbf{h}}} \xrightarrow{a.s.} 0. \tag{7.82}$$

As $\alpha > 1$, (7.21) can be used in (7.82) to obtain

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) - \lim_{\sigma^2\to 0} \frac{\sigma^2\|\bar{\mathbf{h}}\|^2}{\bar{\mathbf{h}}^H\check{\boldsymbol{\Phi}}^{-1}\bar{\mathbf{h}}} \xrightarrow{a.s.} 0. \tag{7.83}$$

Convergence (7.72) follows from (7.83) and the fact that

$$\frac{\|\bar{\mathbf{h}}\|^2}{\bar{\mathbf{h}}^H\check{\boldsymbol{\Phi}}^{-1}\bar{\mathbf{h}}} \leq \lambda_{\max}(\check{\boldsymbol{\Phi}}) \leq \frac{B}{b^2(\alpha-1)}. \tag{7.84}$$

□

From (7.69), it follows that if either α or B goes to zero, $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$ converges to unity. It is an expected fact since either of the above cases implies that the effect of interference is negligible. Moreover, it follows from (7.71) that if $\alpha < 1$ and $b = B$, then $\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = 1 - \alpha$ which, under similar conditions, is equal to $\eta(\mathbf{f}_{\mathrm{mmse}})$ [79]. Finally, similar to the MMSE receiver [79], $\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) = 0$ for $\alpha > 1$.

Note that if, similar to Section 4 of Chapter 6, we use the common for uplink assumption that $h_{k,l}$ are zero-mean circular random variables independent across k and l, then the results of Corollary 2 and Theorem 2 can be significantly simplified. First, note that using (6.111) and (6.112) in (7.19) and (7.23), $\check{\phi}_l$ and $\check{\psi}_l$ can be represented as

$$\check{\phi}_l = \sum_{s=-(L-1)}^{L-1} e^{j\frac{2\pi}{L_c}(l-1)s} \left(\frac{1}{L_c} \sum_{k=2}^{K} \frac{r_{k,s}}{\check{f}_k} \right) \tag{7.85}$$

where

$$\check{f}_k = 1 + \sum_{p=-(L-1)}^{L-1} r_{k,p} \left(\frac{1}{L_c} \sum_{n=1}^{L_c} \frac{e^{j\frac{2\pi}{L_c}(n-1)p}}{\check{\phi}_n} \right) \tag{7.86}$$

and

$$\check{\psi}_l = 1 + \sum_{s=-(L-1)}^{L-1} e^{j\frac{2\pi}{L_c}(l-1)s} \left(\frac{1}{L_c} \sum_{k=2}^{K} \frac{r_{k,s}}{\check{g}_k} \right) \tag{7.87}$$

where

$$\check{g}_k = \sum_{p=-(L-1)}^{L-1} r_{k,p} \left(\frac{1}{L_c} \sum_{n=1}^{L_c} \frac{e^{j\frac{2\pi}{L_c}(n-1)p}}{\check{\psi}_n} \right). \tag{7.88}$$

Similar to our discussion in Section 4 of Chapter 6, it can be concluded from (7.85) and (7.87) that both $\check{f}_k$ and $\check{g}_k$ are the functions of $2L-1$ random variables $r_{k,-(L-1)}, \ldots, r_{k,L-1}$ among which only one is correlated with $r_{k,s}$, and, therefore, both $1/\check{f}_k$ and $1/\check{g}_k$ are approximately uncorrelated with $r_{k,s}$. The above discussion can be used to prove the following theorem.

Theorem 3: Assume that $h_{k,l}$ are zero-mean circular random variables that are independent across k and l and, moreover, $1/\check{f}_k$ and $1/\check{g}_k$ are uncorrelated with $r_{k,s}$ for $s = -(L-1), \ldots, L-1$. Then, for $\alpha > 1$ we have that

$$\check{\mathbf{\Phi}} - \frac{1}{\check{\phi}^*_{\mathbf{h}}} \mathbf{I} \xrightarrow{e.i.p.} \mathbf{0} \tag{7.89}$$

where $\check{\phi}_{\mathbf{h}}^*$ is the unique positive solution to

$$\check{\phi}_{\mathbf{h}} = \frac{1}{L_c}\sum_{k=2}^{K}\frac{\|\mathbf{h}_k\|^2}{1+\|\mathbf{h}_k\|^2/\check{\phi}_{\mathbf{h}}}. \tag{7.90}$$

Moreover, if the empirical distribution of the received user powers $\{\|\mathbf{h}_2\|^2, \|\mathbf{h}_3\|^3, \ldots, \|\mathbf{h}_K\|^2\}$ converges almost surely to the nonrandom distribution $F(p)$, then

$$\check{\phi}_{\mathbf{h}}^* \xrightarrow{a.s.} \gamma^* \tag{7.91}$$

where γ^* is the unique positive solution to

$$\int_0^\infty \frac{p}{\gamma+p}dF(p) = \frac{1}{\alpha}. \tag{7.92}$$

In turn, for $\alpha < 1$,

$$\check{\mathbf{\Psi}} \xrightarrow{e.i.p.} (1-\alpha)\mathbf{I}. \tag{7.93}$$

Proof: Proof of (7.89)-(7.92) is similar to that of Theorem 3 in Chapter 6. To prove (7.93), first note that using a similar argument as in (6.122) and (6.123), it can be shown that

$$\check{\psi}_l - \theta \xrightarrow{i.p.} 0 \tag{7.94}$$

where

$$\theta = 1 + \frac{1}{L_c}\sum_{k=2}^{K}\frac{\|\mathbf{h}_k\|^2}{\check{g}_k}. \tag{7.95}$$

From (7.88) and (7.94), it follows that

$$\check{g}_k - \sum_{p=-(L-1)}^{L-1}\frac{r_{k,p}}{\theta}\left(\frac{1}{L_c}\sum_{n=1}^{L_c}e^{j\frac{2\pi}{L_c}(n-1)p}\right) \xrightarrow{i.p.} 0 \tag{7.96}$$

or, equivalently,

$$\check{g}_k - \frac{\|\mathbf{h}_k\|^2}{\theta} \xrightarrow{i.p.} 0. \tag{7.97}$$

From (7.95) and (7.97) and the fact that $K/L_c \to \alpha$, we obtain

$$\theta \xrightarrow{i.p.} 1/(1-\alpha) \tag{7.98}$$

and, therefore,

$$\check{\psi}_l \xrightarrow{i.p.} \frac{1}{1-\alpha}. \tag{7.99}$$

Using (7.99) in (7.18) yields

$$[\check{\mathbf{\Psi}}]_{1m} \xrightarrow{i.p.} (1-\alpha)\delta_{1m}. \tag{7.100}$$

Equation (7.93) directly follows from (7.100). □

We stress that even if $h_{k,l}$ are statistically independent, $1/\check{f}_k$ and $1/\check{g}_k$ have small correlation with $r_{k,s}$. Hence, in practice, (7.89) and (7.93) do not hold in the exact sense, rather they show that if $h_{k,l}$ are zero-mean circular and independent random variables, then $\check{\mathbf{\Phi}} \approx 1/\check{\phi}^*_{\mathbf{h}}\mathbf{I}$ for $\alpha > 1$ and $\check{\mathbf{\Psi}} \approx (1-\alpha)\mathbf{I}$ for $\alpha < 1$.

From the above discussion along with (7.60) and (7.70), it follows that if the user channel coefficients are zero-mean circular and statistically independent, then we have

$$\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \approx \frac{|\bar{\mathbf{h}}^H \mathbf{h}_1|^2}{\check{\phi}^*_{\mathbf{h}} \|\bar{\mathbf{h}}\|^2}, \qquad \alpha > 1 \tag{7.101}$$

$$\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}})) \approx 1-\alpha, \qquad \alpha < 1. \tag{7.102}$$

Note that the approximate expression of $\eta(\mathbf{f}_{\mathrm{m}}(\bar{\mathbf{h}}))$ in (7.102) implies that if the user channel coefficients are zero-mean circular and independent random variables, then the asymptotic efficiency of the MOE receiver becomes independent from the choice of $\bar{\mathbf{h}}$.

7.2 Simulations

In all examples the length of the user channel vectors $L = 5$ is selected and the entries of the spreading codes are randomly generated from $\pm 1/\sqrt{L_c}$. All figures but the last two are the simulation results of the uplink case.

In Figs. 7.1-7.2, spreading codes of length $L_c = 128$ are randomly assigned to each user and the channel vectors $\mathbf{h}_k$ are randomly and independently generated from a zero-mean circular multivariate Gaussian distribution $\mathcal{N}_C(\mathbf{0}, \frac{\sigma^2_{\mathbf{h}}}{L}\mathbf{I})$.

Fig. 7.1 shows the MSE between the estimated channel vector $\mathbf{h}_{\mathrm{C}}$ and the normalized version of $\mathbf{h}_1$ versus $\sigma^2_{\mathbf{h}}/\sigma^2$ for different values of α. Each simulation point is the average

over 200 realizations of the channel vectors. From Fig. 7.1, it can be observed that, as $\sigma_{\mathbf{h}}^2/\sigma^2$ increases, the MSE consistently decreases for all values of $\alpha < 1$. Interestingly, this observation is no longer true for $\alpha = 1.05$ where the MSE plot shows some estimation error floor. Fig. 7.1 verifies Equation (7.58) which states that $\mathbf{h}_{\mathrm{C}}$ converges to a scaled version of $\mathbf{h}_1$ at high SNRs only if $\alpha < 1$.

Fig. 7.2 shows $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus $\sigma_{\mathbf{h}}^2/\sigma^2$ for the same simulation setup as in Fig. 7.1. It can be observed from Fig. 7.2 that, as $\sigma^2 \to 0$, $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ goes to infinity only if $\alpha < 1$. This further verifies the results of Corollary 1 and Corollary 2.

In Figs. 7.3-7.6, each channel vector $\mathbf{h}_k$ is randomly and independently generated from a zero-mean circular multivariate Gaussian distribution $\mathcal{N}_C(\mathbf{0}, \frac{\sigma_{\mathbf{h}}^2}{L}\mathbf{I})$ and remains fixed during these examples while a total of 200 sets of spreading codes of lengths $L_c = 128$ and $L_c = 256$ are randomly generated for each user and the results are displayed for all sets of spreading codes.

Fig. 7.3 shows $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ versus $\alpha > 1$. In the upper and middle subplots, $\sigma_{\mathbf{h}}^2/\sigma^2$ is selected equal to 10 and 30 dB, respectively, while in the lower subplots $\sigma^2 = 0$ is chosen to simulate the scenario in which $\sigma_{\mathbf{h}}^2/\sigma^2$ converges to ∞. In all subplots, the asymptotic SINR obtained in (6.95) is shown by solid lines while the asymptotic expressions of $\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ obtained in (7.60) and (7.101) are shown by dashed and dash-dot lines, respectively. It can be observed from Fig. 7.3 that the experimental SINRs are very close to the analytical expression of (6.95) for all values of $\sigma_{\mathbf{h}}^2/\sigma^2$ and, moreover, as L_c increases, the spread of the experimental SINRs around the corresponding analytical curves tends to decrease. Note also that the analytical curves of $\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ obtained via (7.60) and (7.101) are almost indistinguishable from each other. This observation further verifies that, if the user channel coefficients are zero mean circular and independent random variables, then (7.101) is a reliable approximation of (7.60). Finally, it can be observed from Fig. 7.3 that, as $\sigma_{\mathbf{h}}^2/\sigma^2$ increases, both the experimental SINRs and the analytical SINR curve of (6.95) converge to the analytical curves of $\lim_{\sigma^2 \to 0} \mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ obtained via (7.60) and (7.101).

The same experiment made for $\bar{\mathbf{h}} = \mathbf{e}_1$ in the last example is also examined for $\bar{\mathbf{h}} = \mathbf{h}_{\mathrm{C}}$

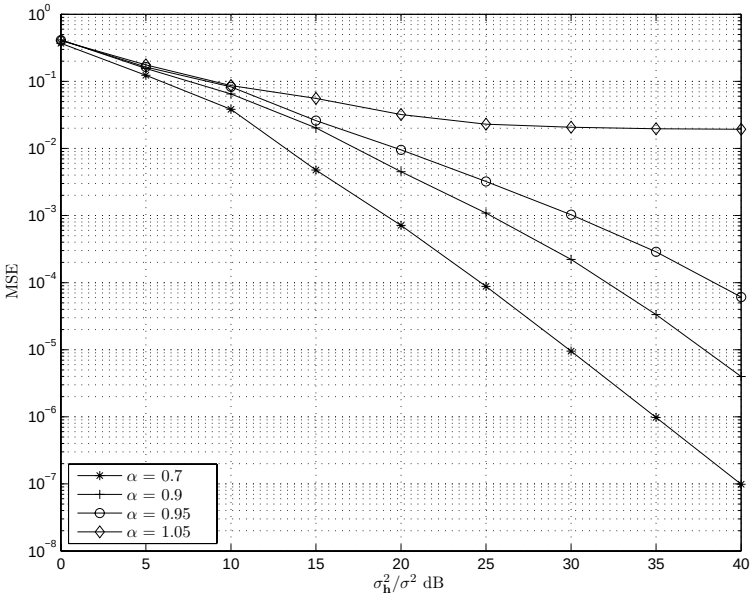

Figure 7.1: MSE of the Capon channel estimate versus $\sigma^2_{\mathbf{h}}/\sigma^2$ for different values of α in the uplink case.

in Fig. 7.4. Similar observations as in Fig. 7.3 can be made from Fig. 7.4 for $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$. This verifies the accuracy of our analytical results for different choices of the constraint vector $\bar{\mathbf{h}}$.

Fig. 7.5 shows the experimental values of $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ versus α. The upper, middle, and lower subplots correspond to $\sigma^2_{\mathbf{h}}/\sigma^2$ being equal to 10, 20, and 40 dB, respectively. The analytical efficiency expression (7.68), as well as the analytical expressions for the asymptotic efficiency $\eta(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ derived in (7.70) and (7.102) are also shown by solid, dashed, and dash-dot lines, respectively. It can be observed from this figure that the experimental and analytical values of $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ are very close over all tested values of α and $\sigma^2_{\mathbf{h}}/\sigma^2$. Moreover, the difference between the experimental and the analytical efficiencies decreases as L_c grows from 128 to 256. Note also that the analytical curves of (7.70) and (7.102) are very close to each other. This verifies that, if $\mathbf{h}_k$ are zero-mean circular and independent random vectors, then (7.102) represents a valid approximation of (7.70). It also can be observed from Fig. 7.5 that, as $\sigma^2_{\mathbf{h}}/\sigma^2$ increases, both the experimental and the analytical $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ become closer to the asymptotic efficiency curves of (7.70) and (7.102). This

verifies that (7.70) and (7.102), in fact, represent $\lim_{\sigma^2 \to 0} \mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$. Finally, it can be observed from the lower subplots that $\eta(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ tends to zero for $\alpha > 1$, confirming our result in (7.72).

Fig. 7.6 displays $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus α, demonstrating that similar results as in Fig. 7.5 also hold true for $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$.

In Fig. 7.7, the spreading factor $L_c = 256$ is selected and a total of 200 sets of spreading codes are randomly generated for each user. This figure shows $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ versus α for all realizations of the spreading codes in the downlink case. A random vector of length $L = 5$ is drawn from a zero-mean complex Gaussian distribution with the covariance matrix

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_1 & \mathbf{0} \\ \mathbf{0} & \gamma_2 \mathbf{I}_4 \end{bmatrix} \tag{7.103}$$

and is used as the channel vector between the base station and the user-of-interest. Transmitted power from the base station to all users is assumed to be equal. In the left subplots, $\vartheta = \gamma_1/\gamma_2 = 0$ dB is chosen, while in the right subplots $\vartheta = 30$ dB is selected to simulate line-of-sight at the first channel coefficient. Moreover, the upper, the middle, and the lower subplots correspond to the scenarios in which $\mathrm{tr}(\mathbf{\Gamma})/\sigma^2$ is equal to 10, 20, and 40 dB, respectively. In addition to the experimental values of $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$, the analytical efficiency expression (7.68) as well as the analytical asymptotic efficiency curves (7.70) and (7.102) are also shown with solid, dashed, and dash-dot lines, respectively. Thick lines are also used to show the lower and the upper bounds on the asymptotic efficiency of (7.71). It can be observed from Fig. 7.7 that the experimental efficiency results are closely followed by the analytical curve of (7.68), and, moreover, as $\mathrm{tr}(\mathbf{\Gamma})/\sigma^2$ increases, both the experimental and the analytical efficiency results converge to the asymptotic efficiency curve of (7.70). Interestingly, Fig. 7.7 shows that (7.70) and (7.102) are not necessarily close, verifying our discussion that (7.102) does not necessarily represent a valid approximation of the asymptotic efficiency in the downlink case. It can also be observed from the lower right subplot that, in the strong line-of-sight case where $\vartheta = 30$ dB, (7.71) provides tight bounds on the asymptotic efficiency. However, as the lower left subplot shows, these bounds become loose in the case of $\vartheta = 0$ dB.

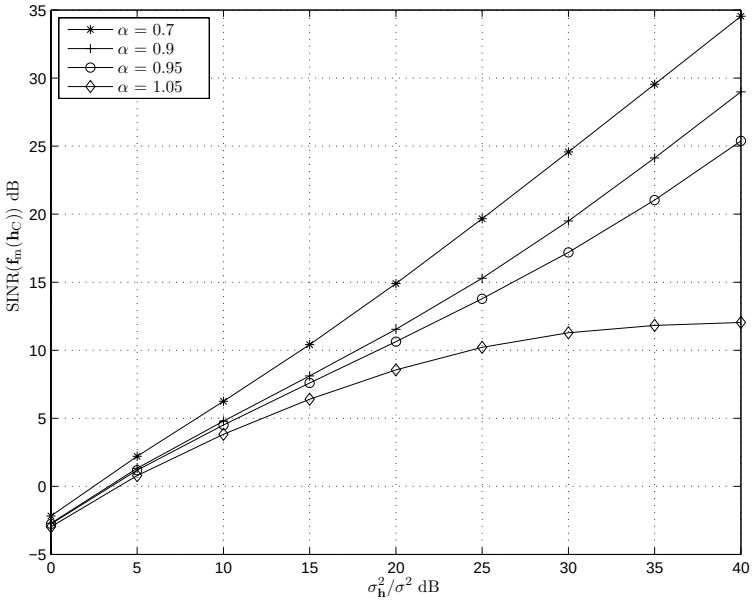

Figure 7.2: $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus $\sigma_{\mathbf{h}}^2/\sigma^2$ for different values of α in the uplink case.

Fig. 7.8 shows that similar observations as in Fig. 7.7 can also be made for the Capon receiver. This verifies that our analytical results are also valid for $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ and $\eta(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ in the downlink case.

7.3 Chapter summary

High SNR properties of the MOE receiver and the Capon channel estimate have been analyzed for the asymptotic scenario that both the spreading factor and the number of users go to infinity with the same rate. It has been shown that, if the system load is less than one, then the Capon channel estimate converges to a scaled version of the channel vector of the user-of-interest. This result has been shown to be consistent with the result obtained in [78] for the nonasymptotic regime, but relaxes the identifiability condition that should hold in the nonasymptotic case. It has also been proved that, if the system load is less than one and the noise power goes to zero, then the SINR of the blind MOE receiver goes to infinity regardless of the interfering user powers.

The asymptotic SINR expression of the MOE receiver (for the case that the system load is larger than one), efficiency, and asymptotic efficiency expressions of the MOE receiver are derived and lower and upper bounds of these expressions are also obtained. Simple approximations of the so-obtained expressions have also been offered for the case that the user channels are zero-mean circular and independent random vectors. In contrary to their nonasymptotic counterparts, all the so-obtained expressions are independent from the user spreading codes and are simple functions of the physical parameters such as the user channel vectors, the constraint vector used in the MOE receiver, and the system load.

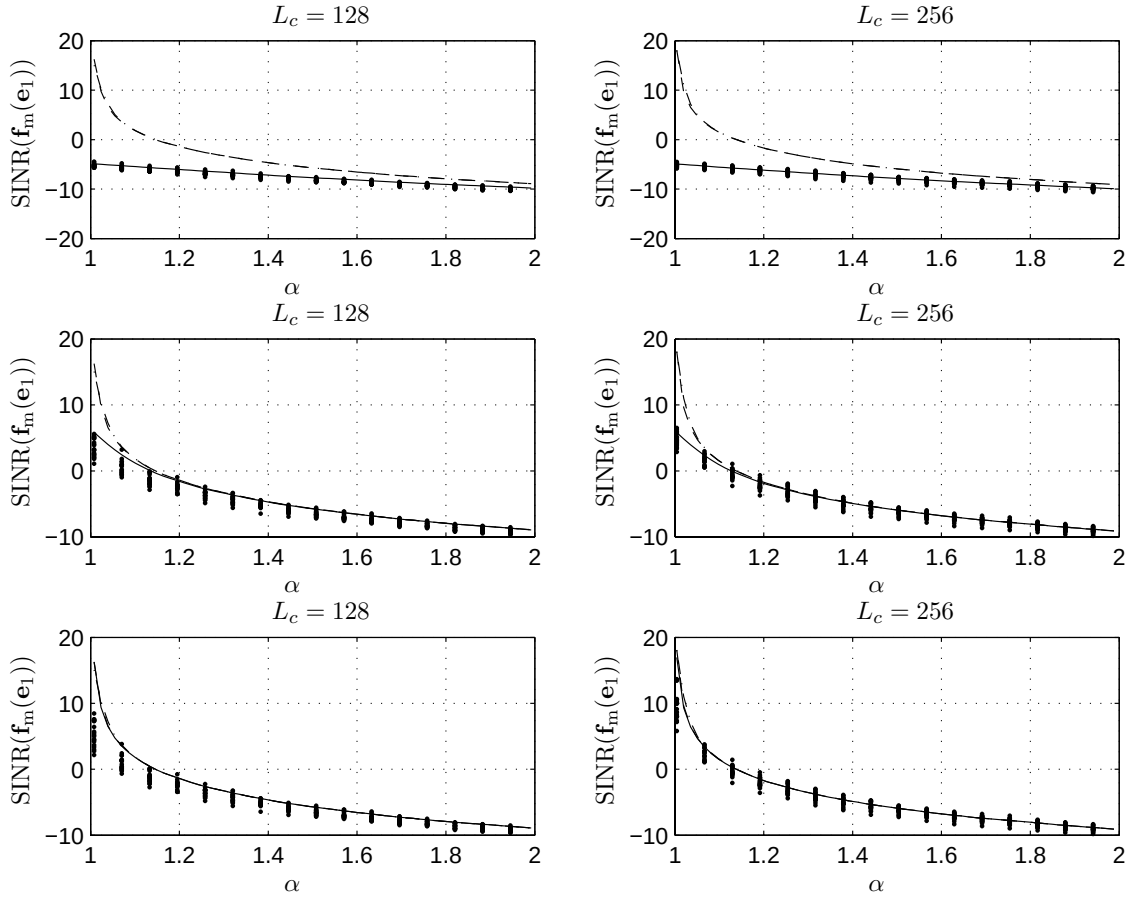

Figure 7.3: $\mathrm{SINR}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ versus $\alpha > 1$ in the uplink case for $\sigma^2_{\mathbf{h}}/\sigma^2 = 10$ dB (upper subplots), $\sigma^2_{\mathbf{h}}/\sigma^2 = 30$ dB (middle subplots), and $\sigma^2_{\mathbf{h}}/\sigma^2 = +\infty$ (lower subplots) in the uplink case.

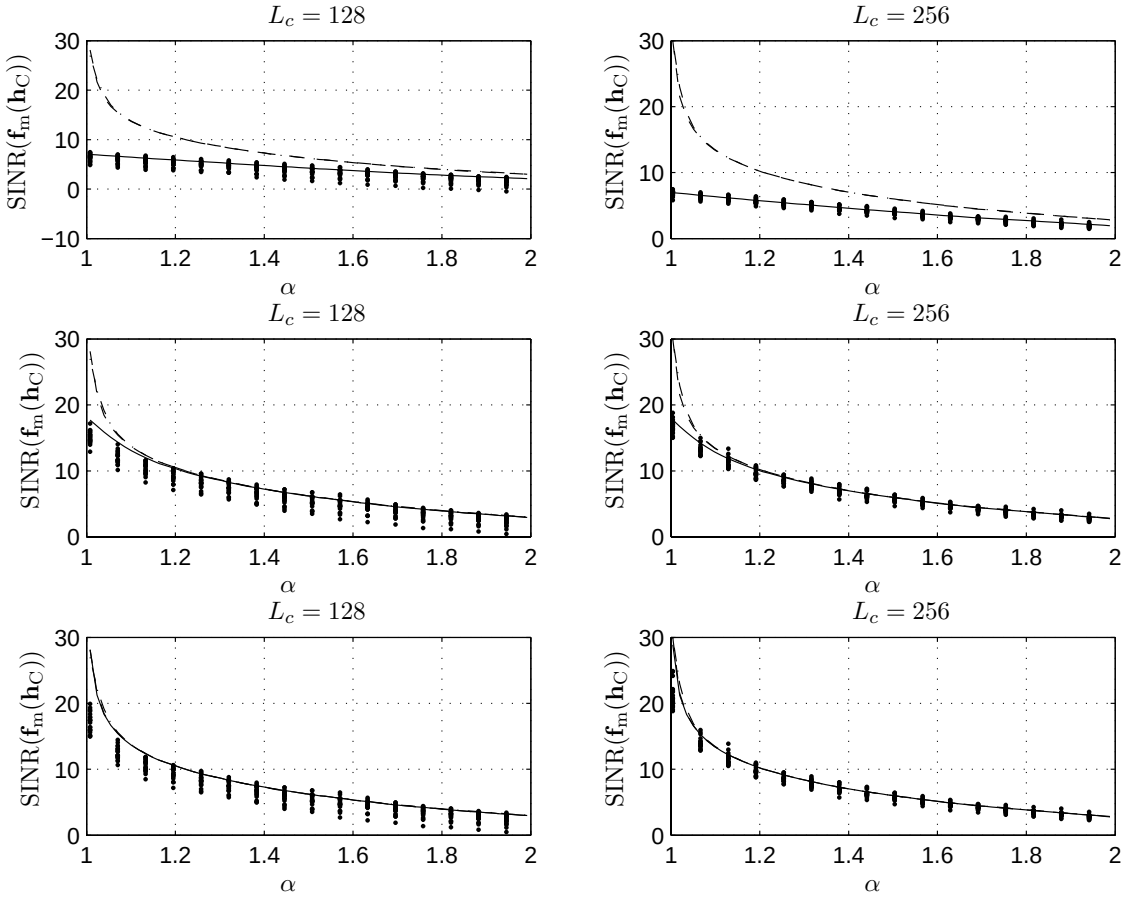

Figure 7.4: $\mathrm{SINR}(\mathbf{f}_\mathrm{m}(\mathbf{h}_\mathrm{C}))$ versus $\alpha > 1$ in the uplink case for $\sigma^2_\mathbf{h}/\sigma^2 = 10$ dB (upper subplots), $\sigma^2_\mathbf{h}/\sigma^2 = 30$ dB (middle subplots), and $\sigma^2_\mathbf{h}/\sigma^2 = +\infty$ (lower subplots) in the uplink case.

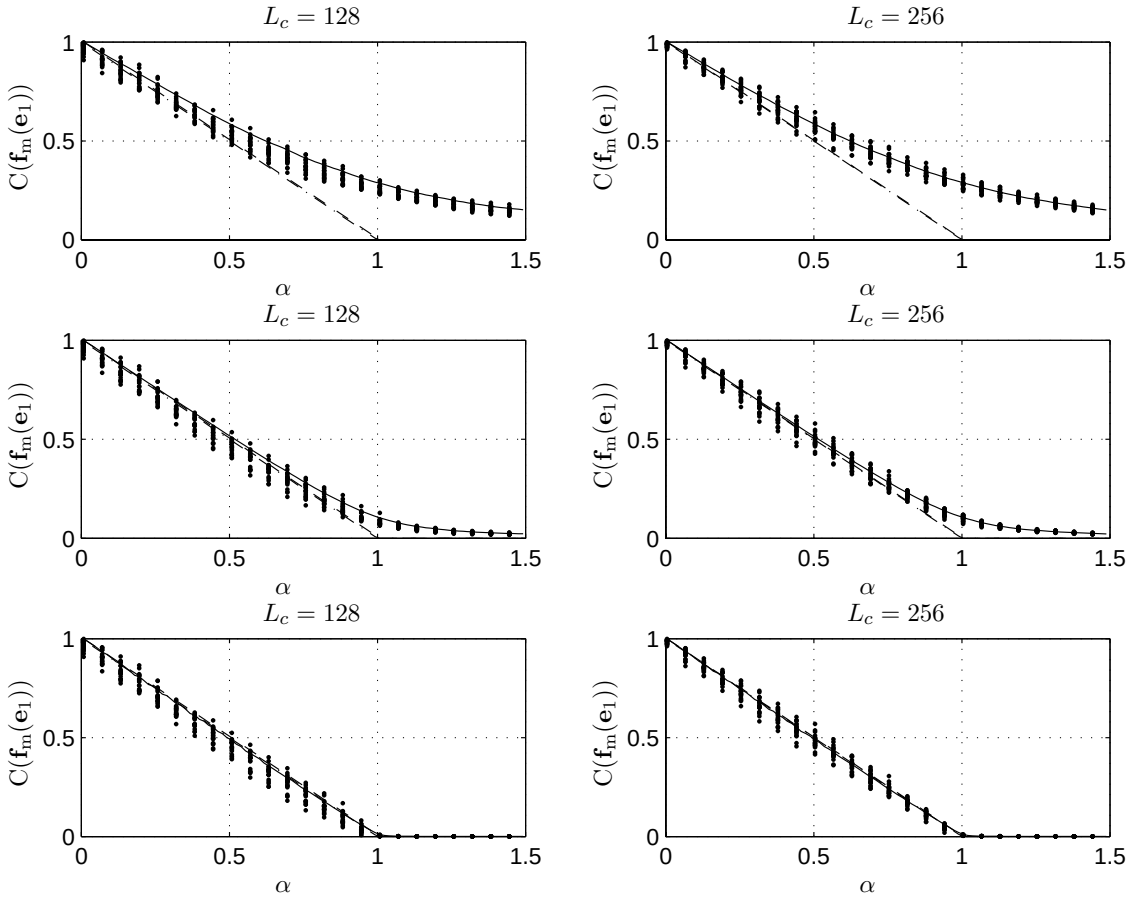

Figure 7.5: $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{e}_1))$ versus α in the uplink case for $\sigma^2_{\mathbf{h}}/\sigma^2 = 10$ dB (upper subplots), $\sigma^2_{\mathbf{h}}/\sigma^2 = 20$ dB (middle subplots), and $\sigma^2_{\mathbf{h}}/\sigma^2 = 40$ dB (lower subplots) in the uplink case.

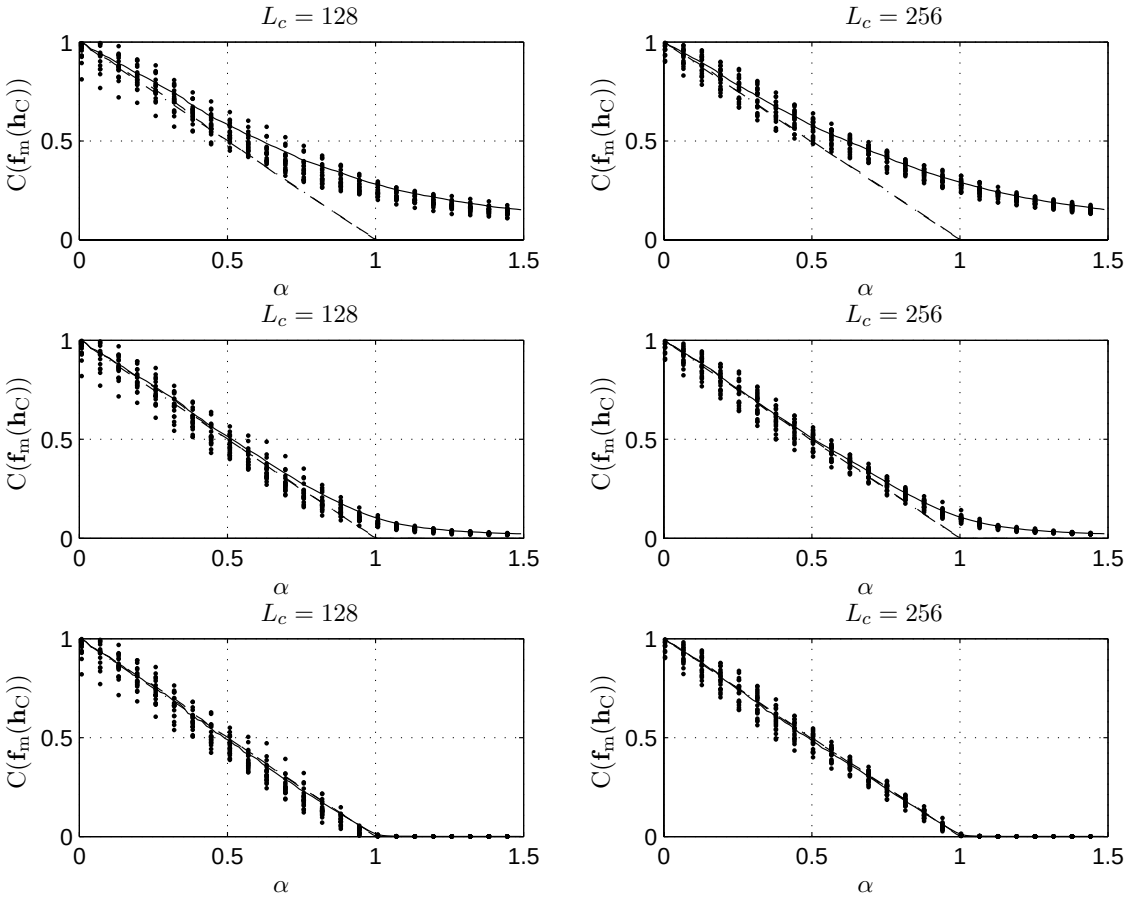

Figure 7.6: $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus α in the uplink case for $\sigma^2_{\mathbf{h}}/\sigma^2 = 10$ dB (upper subplots), $\sigma^2_{\mathbf{h}}/\sigma^2 = 20$ dB (middle subplots), and $\sigma^2_{\mathbf{h}}/\sigma^2 = 40$ dB (lower subplots).

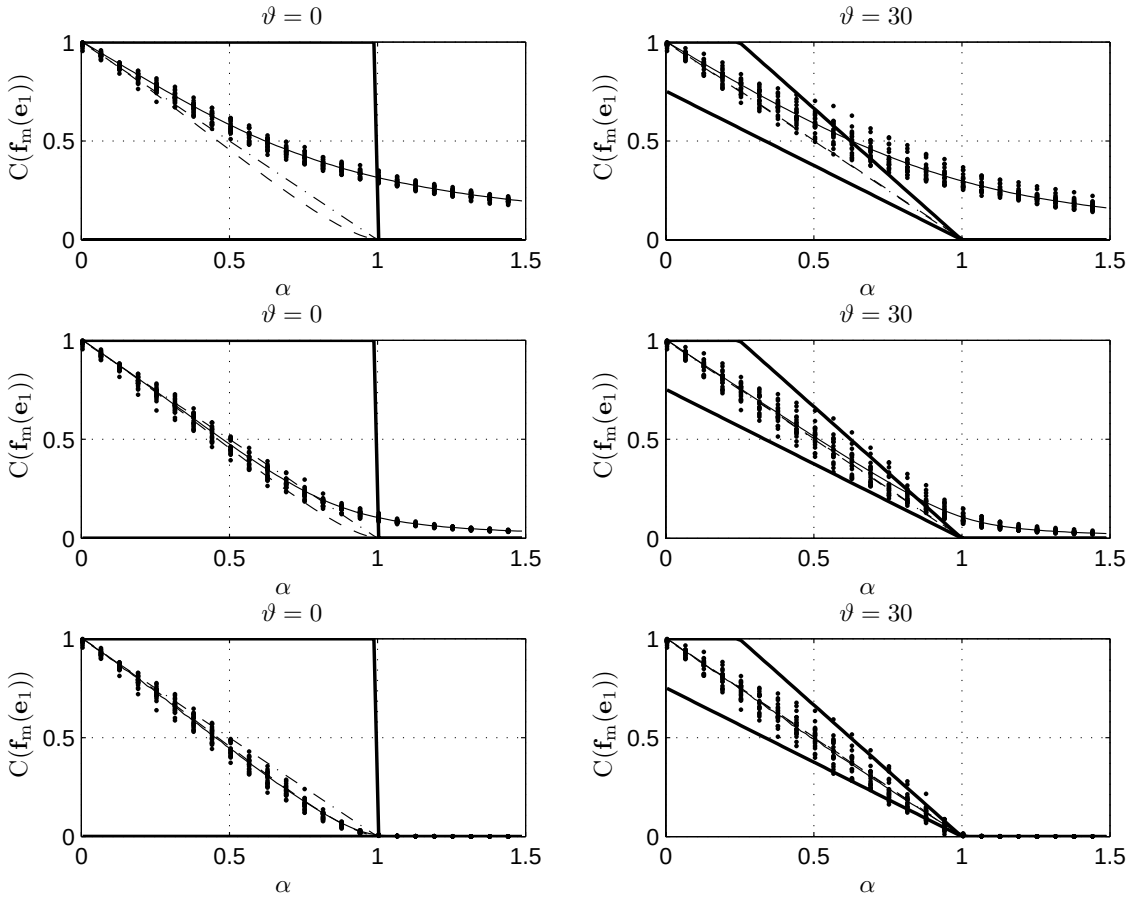

Figure 7.7: $\mathrm{C}(\mathbf{f}_\mathrm{m}(\mathbf{e}_1))$ versus α in the downlink case for $\sigma^2_\mathbf{h}/\sigma^2 = 10$ dB (upper subplots), $\sigma^2_\mathbf{h}/\sigma^2 = 20$ dB (middle subplots), and $\sigma^2_\mathbf{h}/\sigma^2 = 40$ dB (lower subplots).

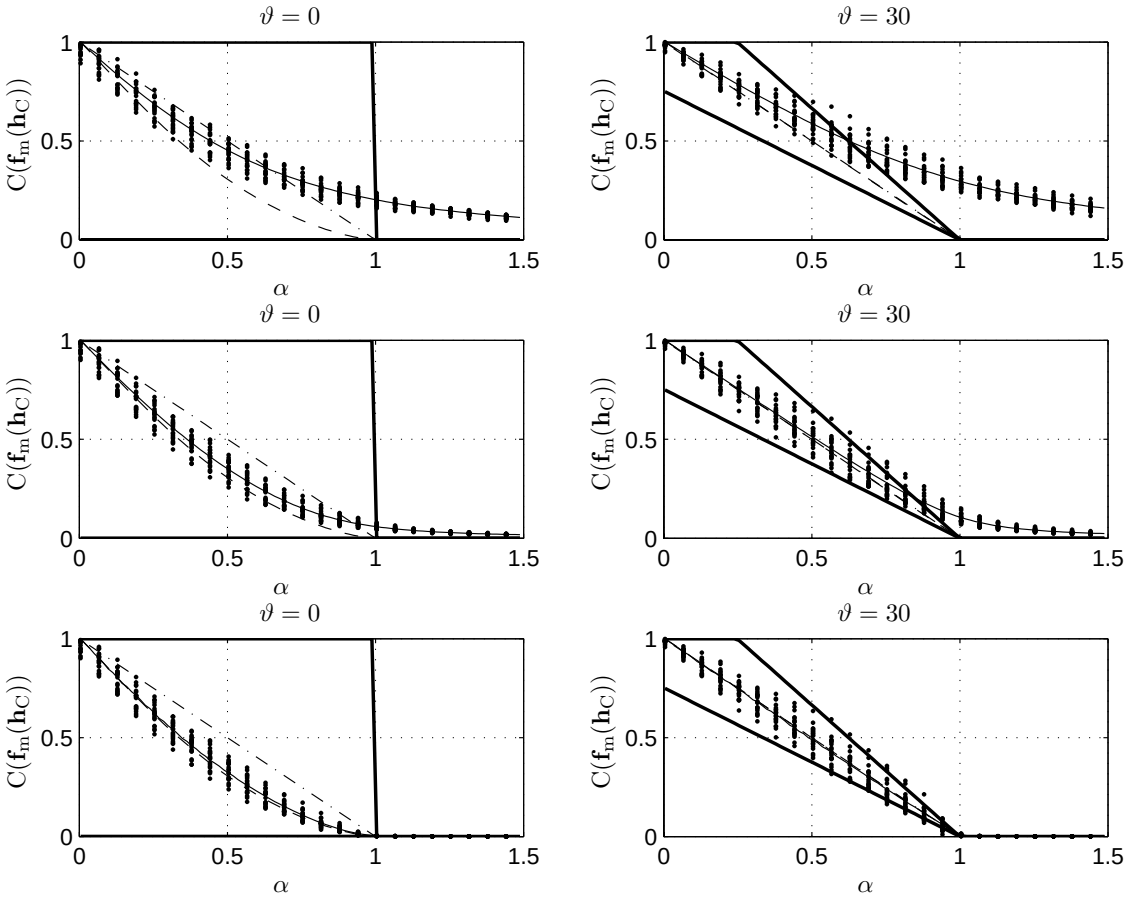

Figure 7.8: $\mathrm{C}(\mathbf{f}_{\mathrm{m}}(\mathbf{h}_{\mathrm{C}}))$ versus α in the downlink case for $\sigma^2_{\mathbf{h}}/\sigma^2 = 10$ dB (upper subplots), $\sigma^2_{\mathbf{h}}/\sigma^2 = 20$ dB (middle subplots), and $\sigma^2_{\mathbf{h}}/\sigma^2 = 40$ dB (lower subplots).

Chapter 8

Conclusions and future work

8.1 Conclusions

In this thesis, we studied and developed several advanced techniques in blind linear signature estimation and multiuser detection for DS-CDMA systems.

In Chapter 2, we considered the problem of signature estimation in the presence of white noise. Assuming that the transmitted signals are drawn from the BPSK constellation, we developed a novel subspace-based estimation technique which, in contrary to the existing subspace-based estimation algorithms, additionally exploits the noncircular property of the transmitted signals along with the circularity of noise to substantially improve the estimation performance. As compared to its predecessors, the proposed technique can estimate longer channel impulse responses and is applicable in scenarios with larger number of active users. The necessary and sufficient identifiability condition of the proposed algorithm has been derived. This condition shows that the proposed technique is able to identify all channel vectors that can be identified using the earlier techniques while the reverse statement is not necessarily true. We considered a practical scenario where the exact data covariance matrix is not available, and the proposed algorithm uses the sample data covariance matrix. For such a scenario, the MSE of the estimated channel vector has been derived and the effects of different parameters on the estimation performance have been discussed. Finally, it

has been shown that the proposed technique can also be easily applied to the oversampled version of the received data to further improve the estimation performance.

The problem of signature estimation in the presence of unknown correlated noise was considered in Chapter 3. In this chapter, the performances of two popular subspace-based signature estimation algorithms for unknown correlated noise have been analyzed. The first technique, known as the WP algorithm [90], assumes that the receiver is equipped with two antennas and noise is spatially white, and, therefore, the signal subspace can be obtained from the data cross-correlation matrix between these two antennas. The second technique, known as the BP algorithm [11], exploits the assumptions that the transmitted signals are BPSK and noise is circular to obtain the signal subspace from the data pseudo-covariance matrix.

In Chapter 3, we have derived the MSE of the estimated channel of the BP and WP algorithms and discussed the effects of various physical parameters on their performance. In particular, we have shown that the performance of these techniques critically depends on the direction of the eigenvectors of the noise covariance matrix. Fixing the eigenvalues of the noise covariance matrix while letting the associated eigenvectors be arbitrary, we obtained both the maximum and the minimum possible MSEs and the corresponding sets of eigenvectors which result in such extreme MSE values. We also obtained the extreme MSE values over all noise covariance matrices with identical average noise power and have shown that both the maximum and the minimum MSEs are achieved when the noise covariance matrix is rank-deficient. As a rank-deficient noise covariance matrix is usually the result of narrow-band interference, we have concluded that both WP and BP algorithms are quite sensitive to narrow-band interferering signals.

We also showed that the MSE performances of both the WP and BP algorithms in the presence of white noise are identical to that of the conventional LX algorithm which its application is limited to the white noise case. Finally, we proved that the performance of the WP algorithm in the high SNR regime becomes independent from the noise covariance matrix and the user received power at the second auxiliary antenna. Generalizing this property to the multiple antenna systems, we have shown that the choice of the auxiliary

antenna is arbitrary at high SNRs.

In Chapter 4 we continued our study of the blind signature estimation problem in the presence of unknown correlated noise. In this chapter, we have proposed a novel signature estimation technique which, in contrast to the WP algorithm, does not require any second antenna, and, moreover, unlike the BP algorithm, can be applied to an arbitrary signal constellation. The proposed technique only requires the unknown correlated noise to be in wide-sense stationary; the assumption which has been widely used in the literature. We have exploited the centro-Hermitian property of the noise covariance matrix along with the idea of covariance differencing [63] to obtain a low-rank matrix which only depends on the user signals. Applying a standard subspace-based estimation technique to the latter matrix, we have obtained the resulting channel vector estimate. We have also derived the necessary and sufficient channel identifiability condition and have obtained the MSE of the channel vector estimate. We have also shown that the proposed algorithm can be applied to temporally oversampled received data to improve the estimation performance.

In Chapter 5, we have developed a novel blind linear multiuser receiver that is robust against the effects of the signature estimation error as well as the error resulting from estimating the data covariance matrix. Assuming that the norms of the signature and the data covariance matrix estimation errors are bounded by known values, we have obtained the maximum of the MSE cost function and, then, derived the robust receiver vector that minimizes the so-obtained worst-case MSE. We have shown that our robust receiver is, in fact, the standard MMSE receiver that uses the sample data covariance matrix with an adaptive diagonal loading. It has been proven that the optimal diagonal loading value can be found as the unique solution to a certain nonlinear equation and an efficient technique to solve this equation has also been proposed. Simulation results have demonstrated that the proposed robust receiver can achieve substantial performance improvements as compared to its conventional non-robust receiver counterparts, while enjoying the similar computational complexity.

In Chapter 6, we have analyzed the asymptotic performance of the MOE receiver in a large random DS-CDMA system where the user spreading codes are i.i.d. random vectors

and both the spreading factor and the number of users go to infinity with the same rate. The asymptotic properties of the optimized version of the MOE receiver (commonly referred to as the Capon receiver) and the Capon channel estimate have also been investigated. In this chapter, we first have assumed that the delayed versions of the spreading codes are statistically independent random vectors and have obtained approximate SINR expressions of the MOE and the Capon receivers. Under the above assumption, it has been shown that the SINR of the Capon receiver converges to that of the optimal MMSE receiver, and the Capon channel estimate converges to a scaled version of the channel vector of the user-of-interest.

We have then dropped the simplifying assumption that the delayed versions of the spreading codes are statistically independent, and have obtained the asymptotic SINR expressions of the MOE and the Capon receivers for the circular Gaussian spreading code case. We have proved that if the user channels are either zero-mean and independent random vectors or single-path, then the results obtained based on the statistical independence of the delayed versions of the spreading codes are valid approximations. As the assumption of independent user channels is valid in the uplink case, while is irrelevant in the downlink case, it has been concluded that the performance of the Capon receiver remains very close to that of the optimum MMSE receiver in the uplink case while the performances of these two receivers can be considerably different in the downlink case.

Computer simulations have verified our analytical results and have shown that the asymptotic results obtained for the circular Gaussian spreading codes are also approximately valid for more general i.i.d. spreading codes.

In Chapter 7, we have used the results of Chapter 6 to analyze the high SNR properties of the MOE receiver and the Capon channel estimate in large random DS-CDMA systems. It has been shown that, if the system load is less than one, then the Capon channel estimate converges to a scaled version of the actual channel vector. This result relaxes the identifiability condition obtained in [78]. Moreover, as SNR goes to infinity, it has been proved that the SINR of the MOE receiver goes to infinity provided that the system load is less than one while this quantity is upper-bounded in the case when the system load is larger than

one.

As SNR goes to infinity, the asymptotic SINR expressions (corresponding to the case when the system load is larger than one) and the asymptotic efficiency of the MOE receiver have been analyzed and tight lower and upper bounds on these two performance measures have been derived. In the case when the user channel vectors are zero-mean independent random vectors, simplified versions of the above expressions have also been obtained. All the so-obtained asymptotic expressions are independent from the user spreading codes and explicitly explain the asymptotic properties of the MOE receiver and the Capon channel estimate in terms of the physical parameters such as the user channel vectors, the constraint vector used in the MOE receiver, and the system load.

8.2 Future work

An interesting extension of this thesis would be to generalize the algorithm of Chapter 2 to any non-circular signal constellation other than BPSK such as ASK and OQAM, as well as to the CDMA systems which employ QPSK constellation with phase and quadrature components modulated by different spreading sequences. Note that the latter scheme has been adopted in some modes of UMTS [81].

As discussed in Chapters 2-4, the subspace-based estimation techniques are usually based on the eigendecomposition or SVD of (the sample estimate of) the second-order statistics of the received data. However, it has been shown in [97] that these techniques may be implemented using a more computationally complex decomposition technique called CCD [4], [97]. It has been proved for a particular subspace-based DOA estimation technique that if CCD is used to estimate the signal and noise subspaces instead of SVD, then the resulting estimated DOA is more accurate [96]. Motivated by this fact but without any performance analysis, CCD-based versions of the WP and BP algorithms have also been presented in [90] and [11], respectively. An interesting extension of the work in Chapter 3 would be the performance analysis of the CCD-based WP and BP algorithms in comparison to the SVD-based WP and BP algorithms.

It can also be shown that, at the cost of an increased computational complexity, the CCD-based version of the algorithm proposed in Chapter 4 can further improve the signature estimation performance. This is the subject of an ongoing research project of the thesis author.

The robust multiuser receiver proposed in Chapter 5 can be further improved in the following way. Note from (1.12) that the signature vector has a particular structure that can be incorporated in the receiver to improve the detection performance. Moreover, our robust receiver can be extended to multi-antenna systems. In such a case, depending on whether the channel impulse responses between the user-of-interest and the receive antennas are independent or correlated, different robust receivers can be obtained.

Using advanced techniques from random matrix theory, it can be analytically investigated whether the results of Chapters 6-7 are valid for more general random spreading codes than the Gaussian ones. The results of Chapters 6-7 can be extended to asynchronous DS-CDMA systems. Finally, the results of these chapters can be extended to the MOE receivers in space-time coded multi-antenna DS-CDMA systems.

Bibliography

[1] Abramovich Y. I. and Nevrev A. I., "An analysis of effectiveness of adaptive maximization of the signal-to-noise ratio which utilizes the inversion of the estimated correlation matrix," *Radio Engineering and Electronic Physics*, vol. 26, pp. 67-74, Dec. 1981.

[2] Aktas E. and Mitra U., "Single-user sparse channel acquisition in multiuser DS-CDMA systems," *IEEE Transactions on Communications*, vol. 51, pp. 682-693, Apr. 2003.

[3] Amin M. G., Wang C., and Lindsey A. R., "Optimum interference excision in spread spectrum communications using open-loop adaptive filters," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1966-1976, July 1999.

[4] Anderson T. W., *An Introduction to Multivariate Statistical Analysis.* John Wiley & Sons, Second Edition, New York, NY, 1984.

[5] Bai Z. D. and Yin Y. Q., "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," *Annals of Probability*, vol. 21, no. 3, pp. 1275-1294, 1993.

[6] Bensley S. E. and Aazhang B., "Subspace-based channel estimation for code-division multiple-access communication systems," *IEEE Transactions on Communications*, vol. 44, pp. 1009-1020, Aug. 1996.

[7] Buzzi S. and Lops M., "Performance analysis for the improved linear multiuser detectors in BPSK-modulated DS-CDMA systems," *IEEE Transactions on Communications*, vol. 51, pp. 37-42, Jan. 2003.

[8] Buzzi S., Lops M., and Poor H. V., "Code-aided interference suppression for DS/CDMA overlay systems," in *Proc. IEEE,* vol. 90, pp. 394-435, March 2002.

[9] Buzzi S., Lops M., and Tulino A. M., "A generalized minimum-mean-output-energy strategy for CDMA signals with improper MAI," *IEEE Transactions on Information Theory*, vol. 48, pp. 761-767, March 2002.

[10] Buzzi S., Lops M., and Tulino A. M., "A new family of MMSE multiuser receivers for interference suppression in DS/CDMA systems employing BPSK modulation," *IEEE Transactions on Communications*, vol. 49, pp. 154-167, Jan. 2001.

[11] Buzzi S. and Poor H. V., "A single-antenna blind receiver for multiuser detection in unknown correlated noise," *IEEE Transactions on Vehicular Technology*, vol. 51, pp. 209-215, Jan. 2002.

[12] Capon J.,"High-resolution frequency-wavenumber spectrum analysis," in *Proc. IEEE*, vol. 57, pp. 1408-1418, 1969.

[13] Charge P., Wang Y., and Saillard J., "An extended cyclic MUSIC algorithm," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1695-1701, July 2003.

[14] Chaufray J.-M., Hachem W., and Loubaton P., "Asymptotic analysis of optimum and suboptimum CDMA downlink MMSE receivers," *IEEE Transactions on Information Theory*, vol. 50, pp. 2620-2638, Nov. 2004.

[15] Coope I. D. and Renaud P. F., "Trace inequalities with applications to orthogonal regression and matrix nearness problems," Department of Mathematics and Statistics, University of Canterbury, Report UCDMS2000/17 Nov. 2000.

Online:http://www.math.canterbury.ac.nz/research/ucdms2000n17.pdf

[16] Cui S., Luo Z.-Q., and Ding Z., "Robust blind multiuser detection against CDMA signature mismatch," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Salt Lake City, UT, USA, May 2001, pp. 2297-2300.

[17] Debbah M., Hachem W., Loubaton P., and M. de Courville M., "MMSE analysis of certain large isometric random precoded systems," *IEEE Transactions on Information Theory*, vol. 49, pp. 1293-1311, May 2003.

[18] Debbah M. and Müller R. R., "MIMO channel modeling and principle of maximum entropy," *IEEE Transactions on Information Theory*, vol. 51, pp. 1667-1690, May 2005.

[19] Evans J. and Tse D. N. C., "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, pp. 2059-2078, Sep. 2000.

[20] Fu Y. and Leung H., "Narrow-band interference cancellation in spread-spectrum communication systems using chaos," *IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications*, vol. 48, pp. 847-858, July 2001.

[21] Garth L. M. and Poor H. V., "Narrowband interference suppression in impulsive channels," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, pp. 15-34, Jan. 1992.

[22] Gelli G., Paura L., and Ragozini A. R. P., "Blind widely linear multiuser detection," *IEEE Communications Letters*, vol. 4, pp. 187-189, June 2000.

[23] Gershman A. B. and Sidiropoulos N. D., *Space-Time Processing for MIMO Communications.* John Wiley & Sons, West Sussex, England, 2005.

[24] Gerstacker W. H., Schober R., and Lampe A., "Receivers with widely linear processing for frequency-selective channels," *IEEE Transactions on Communications*, vol. 51, pp. 1512-1523, Sep. 2003.

[25] Giannakis G. B., Wang Z., Scaglione A., and Barbarossa S., "AMOUR–Generalized multicarrier transceivers for blind CDMA regardless of multipath," *IEEE Transactions on Communications*, vol. 48, pp. 2064-2076, Dec. 2000.

[26] Girko V. L., *Theory of Stochastic Canonical Equations.* Kluwer Academic Publishers, the Netherlands, 2001.

[27] Haardt M. and Römer F., "Enhancement of unitary ESPRIT for non-circular sources," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, May 2004, pp. 101-104.

[28] Hachem W., "Simple polynomial detectors for CDMA downlink transmissions on frequency-selective channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 164-171, Jan. 2004.

[29] Hanley S. V. and Tse D. N. C., "Resource pooling and effective bandwidths in CDMA networks with multiuser receivers and spatial diversity," *IEEE Transactions on Information Theory*, vol. 47, pp. 1328-1351, May 2001.

[30] Haykin S., *Adaptive filter theory.* Prentice Hall Information and System Sciences Series, Fourth Edition, Upper Saddle River, NJ, 2002.

[31] Honig M., Madhow U., and Verdú S., "Blind adaptive multiuser detection," *IEEE Transactions on Information Theory,* vol. 41, pp. 944-960, July 1995.

[32] Honig M. and Tsatsanis M. K., "Adaptive techniques for multiuser CDMA receivers," *IEEE Signal Processing Magazine*, vol. 17, pp. 49-61, May 2000.

[33] Honig M. and Xiao W., "Performance of reduced-rank linear interference suppression," *IEEE Transactions on Information Theory*, vol. 47, pp. 1928-1946, July 2001.

[34] Horn R. A. and Johnson C. R., *Matrix Analysis.* Cambridge University Press, Cambridge, UK, 1999.

[35] Kiran G. and Tse D. N. C., "Effective interference and effective bandwidth of linear multiuser receivers in asynchronous CDMA systems," *IEEE Transactions on Information Theory*, vol. 46, pp. 1426-1447, July 2000.

[36] Klein A., Kaleh G. K., and Baier P. W., "Zero forcing and minimum mean-square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 276-287, May 1996.

[37] Klein A., Steiner B., and Steil A., "Known and novel diversity approaches as a powerful means to enhance the performance of cellular mobile radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 14, pp. 1784-1795, Dec 1996.

[38] Kondo S. and Milstein B., "Performance of multicarrier DS CDMA systems," *IEEE Transactions on Communications*, vol. 44, pp. 238-246, Feb. 1996.

[39] Li C., Hu G., and Liu M., "Narrow-band interference excision in spread-spectrum systems using self-orthogonalizing transform-domain adaptive filters," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 403-406, March 2000.

[40] Li F., Liu H., and Vaccaro R. J., "Performance analysis for DOA estimation algorithms: unification, simplification, and observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, pp. 1170-1184, Oct. 1993.

[41] Li H., "Blind channel estimation for multicarrier systems with narrowband interference suppression," *IEEE Communication Letters*, vol. 7, pp. 326-328, July 2003.

[42] Li H., Lu X., and Giannakis G. B., "Capon multiuser receiver for CDMA systems with space-time coding," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1193-1204, May 2002.

[43] Li J., Stoica P., and Wang Z., "On robust capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1702-1715, July 2003.

[44] Li L., Tulino A. M., and Verdú S., "Design of reduced-rank MMSE multiuser detectors using random matrix methods," *IEEE Transactions on Information Theory*, vol. 50, pp. 986-1008, June 2004.

[45] Li L.-M. and Milstein L. B., "Rejection of narrow-band interference in PN spread spectrum systems using transversal filters," *IEEE Transactions on Communications*, vol. 30, pp. 925-928, May 1982.

[46] Li Q., Georghiades C. N., and Wang X., "Blind multiuser detection in uplink CDMA with multipath fading: A sequential EM approach," *IEEE Transactions on Communications*, vol. 52, pp. 71-81, Jan. 2004.

[47] Liu H. and Xu G., "A subspace method for signature waveform estimation in synchronous CDMA systems," *IEEE Transactions on Communications*, vol. 44, pp. 1346-1354, Oct. 1996.

[48] Lops M., Ricci G., and Tulino A. M., "Narrow-band-interference suppression in multiuser CDMA systems," *IEEE Transactions on Communications*, vol. 46, pp. 1163-1175, Sep. 1998.

[49] Lupas R. and Verdú S., "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 35, pp. 123-136, Jan. 1989.

[50] Madhow U. and Honig M. L., "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Transactions on Communications*, vol. 42, pp. 3178-3188, Dec. 1994.

[51] Masry E., "Closed-form analytical results for the rejection of narrow-band interference in PN spread-spectrum systems–Part I: Linear prediction filters," *IEEE Transactions on Communications*, vol. 32, pp. 888-896, Aug. 1984.

[52] Mestre X., Fonollosa J. R., and Pagès-Zamora A., "Capacity of MIMO channels: Asymptotic evaluation under correlated fading," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 829-838, June 2003.

[53] Miller S. L., Honig M. L., and Milstein L. B., "Performance analysis of MMSE receivers for DS-CDMA in frequency-selective fading channels," *IEEE Transactions on Communications*, vol. 48, pp. 1919-1929, Nov. 2000.

[54] Mitra U. and Poor H. V., "Adaptive receiver algorithms for near-far resistant CDMA," *IEEE Transactions on Communications*, vol. 43, pp. 1713-1724, Feb.-Apr. 1995.

[55] Möller R. R., "A random matrix model of communication via antenna arrays," *IEEE Transactions on Information Theory*, vol. 48, pp. 2495-2506, Sep. 2002.

[56] Neeser F. D. and Massey J. L., "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, pp. 1293-1302, July 1993.

[57] Pham D. S. and Zoubir A. M., "A sequential algorithm for robust parameter estimation," *IEEE Signal Processing Letters*, vol. 12, pp. 21-24, Jan. 2005.

[58] Peng M. and Wang W., "A framework for investigating radio resource management algorithms in TD-SCDMA systems," *IEEE Radio Communications Magazine*, vol. 43, pp. 12-18, June 2005.

[59] Phoel W. G. and Honig M. L., "Performance of coded DS-CDMA with pilot-assisted channel estimation and linear interference suppression," *IEEE Transactions on Communications*, vol. 50, pp. 822-832, May 2002.

[60] Picinbono B., "On circularity," *IEEE Transactions on Signal Processing*, vol. 42, pp. 3473-3482, Dec. 1994.

[61] Poor H. V., "Active interference suppression in CDMA overlay systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 4-20, Jan. 2001.

[62] Poor H. V. and Wang X., "Code-aided interference suppression for DS/CDMA communications–Part I: Interference suppression capability," *IEEE Transactions on Communications*, vol. 45, pp. 1101-1111, Sep. 1997.

[63] Prasad S., Williams R. T., Mahalanabis A. K., and Sibul L. H.,"A transform-based covariance matrix differentiating approach for some classes of parameter estimation problems," *IEEE Transactions on Acoustics, Speech, Signal Processing*, vol. 36, pp. 631-641, May 1988.

[64] Reynolds D., Wang X., and Poor H. V., "Blind adaptive space-time multiuser detection with multiple transmitter and receiver antennas," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1261-1276, June 2002.

[65] Rong Y., Shahbazpanahi S., and Gershman A. B., "Robust linear receivers for space-time block coded multiaccess MIMO systems with imperfect channel state information," *IEEE Transactions on Signal Processing*, vol. 51, pp. 3081-3090, Aug. 2005.

[66] Roy S. and Li C., "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 572-579, Oct. 2002.

[67] Rusch L. A. and Poor H. V., "Narrowband interference suppression in CDMA spread spectrum communications," *IEEE Transactions on Communications*, vol. 42, pp. 1969-1979, Feb.-Apr. 1994.

[68] Schober R., Gerstacker W. H., and Lampe L. H.-J., "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Transactions on Signal Processing*, vol. 52, pp. 746-756, March 2004.

[69] Schreier P. J. and Scharf L. L., "Second-order analysis of improper complex random vectors and processes," *IEEE Transactions on Signal Processing*, vol. 51, pp. 714-725, March 2003.

[70] Shahbazpanahi S. and Gershman A. B., "Robust blind multiuser detection for synchronous CDMA systems using worst-case performance optimization," *IEEE Transactions on Wireless Communications,* vol. 3, pp. 2232-2245, Nov. 2004.

[71] Shahbazpanahi S., Gershman A. B., Luo Z.-Q., and Wong K. M., "Robust adaptive beamforming for general-rank signal models," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2257-2269. Sep. 2003.

[72] Sidiropoulos N. D., Giannakis G. B., and Bro R., "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, pp. 810-823, March 2000.

[73] Sourour E. A. and Nakagawa M., "Performance of orthogonal multicarrier CDMA in a multipath fading channel," *IEEE Transactions on Communications*, vol. 44, pp. 356-367, March 1996.

[74] Theodoridis S., Kalouptsidis N., Proakis J., and Koyas G., "Interference rejection in PN spread-spectrum systems with LS linear phase FIR filters," *IEEE Transactions on Communications*, vol. 37, pp. 991-994, Sep. 1989.

[75] Tian Z., Bell K. L., and Van Trees H. L., "Robust constrained linear receivers for CDMA wireless systems," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1510-1522, July 2001.

[76] Torlak M. and Xu G., "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Transactions on Signal Processing*, vol. 45, pp. 137-147, Jan. 1997.

[77] Tsatsanis M. K. and Giannakis G. B., "Optimal decorrelating receivers for DS-CDMA systems: a signal processing framework," *IEEE Transactions on Signal Processing*, vol. 44, pp. 3044-3055, Dec. 1996.

[78] Tsatsanis M. K. and Xu Z., "Performance analysis of minimum variance CDMA receivers," *IEEE Transactions on Signal Processing*, vol. 46, pp. 3014-3022, Nov. 1998.

[79] Tse D. N. C. and Hanly S. V., "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, pp. 641-657, March 1999.

[80] Tulino A. M., Lozano A., and Verdú S., "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Transactions on Information Theory*, vol. 51, pp. 2491-2509, July 2005.

[81] Tulino A. M. and Verdú S., "Asymptotic analysis of improved linear receivers for BPSK-CDMA subject to fading," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1544-1555, Aug. 2001.

[82] Ugweje O. C. and Matin S. A. "Performance of CDMA overlay systems using filtering and diversity," *IEEE Transactions on Vehicular Technology*, vol. 52, pp. 456-462, March 2003.

[83] Van Trees H. L., *Optimum Array Processing*, John Wiley & Sons, New York, NY, 2002.

[84] Verdú S., *Multiuser Detection.* Cambridge University Press, Cambridge, UK, 1998.

[85] Verdú S. and Shamai S., "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 45, pp. 622-640, March 1999.

[86] Vorobyov S. A., Gershman A. B., and Luo Z.-Q., "Robust adaptive beamforming using worst-case performance optimization via second-order cone programming," *IEEE Transactions on Signal Processing*, vol. 51, pp. 313-324, Feb. 2003.

[87] Vorobyov S. A., Rong Y., Sidiropoulos N. D., and Gershman A. B., "Robust iterative fitting of multilinear models," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2678-2689, Aug. 2005.

[88] Wang K.-J., Zhou Z.-C., and Yao Y., "Performance analysis of nonlinear interpolation filters in DS spread spectrum systems under narrowband interference condition," *Electronic Letters*, vol. 34, pp. 1464-1465, July 1998.

[89] Wang X. and Poor H. V., "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Transactions on Communications*, vol. 46, pp. 91-103, Jan. 1998.

[90] Wang X. and Poor H. V., "Blind joint equalization and multiuser detection for DS-CDMA in unknown correlated noise," *IEEE Transactions on Circuits and Systems–II: Analog and Digital Signal Processing*, vol. 46, pp. 886-895, July 1999.

[91] Wang X. and Poor H. V., "Blind multiuser detection: A subspace approach," *IEEE Transactions on Information Theory*, vol. 44, pp. 677-690, March 1998.

[92] Wang X. and Poor H. V., "Robust multiuser detection in non-Gaussian channels," *IEEE Transactions on Signal Processing*, vol. 47, pp. 289-305, Feb. 1999.

[93] Wang X. and Poor H. V., "Space-time multiuser detection in multipath CDMA channels," *IEEE Transactions on Signal Processing*, vol. 47, pp. 2356-2374, Sep. 1999.

[94] Wang Z. and Giannakis G. B., "Block precoding for MUI/ISI-resilient generalized multicarrier CDMA with multirate capabilities," *IEEE Transactions on Communications*, vol. 49, pp. 2016-2027, Nov. 2001.

[95] Wilkinson J. H., *The Algebraic Eigenvalue Problem.* Oxford University Press, Oxford, UK, 1999.

[96] Wu Q. and Wong K. M., "Estimation of DOA in unknown noise: Performance analysis of UN-MUSIC and UN-CLE, and the optimality of CCD," *IEEE Transactions on Signal Processing*, vol. 43, pp. 454-468, Feb. 1995.

[97] Wu Q. and Wong K. M., "UN-MUSIC and UN-CLE: An application of generalized correlation analysis to the estimation of the direction of arrival of signals in unknown correlated noise," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2331-2343, Sep. 1994.

[98] Wu S.-H., Mitra U., and Kuo C.-C. J, "Reduced-rank multistage receivers for DS-CDMA in frequency-selective fading channels," *IEEE Transactions on Communications*, vol. 53, pp. 366-378, Feb. 2005.

[99] Xu Z., "On the second-order statistics of the weighted sample covariance matrix," *IEEE Transactions on Signal Processing*, vol. 51, pp. 527-534, Feb. 2003.

[100] Xu Z. and Liu P., "Blind multiuser detection by kurtosis maximization/minimization," *IEEE Signal Processing Letters*, vol. 11, pp. 1-4, Jan. 2004.

[101] Xu Z., Liu P., and Wang X., "Blind multiuser detection: From MOE to subspace methods," *IEEE Transactions on Signal Processing*, vol. 52, pp. 510-524, Feb. 2004.

[102] Xu Z. and Tsatsanis M. K., "Blind adaptive algorithms for minimum variance CDMA receivers," *IEEE Transactions on Communications*, vol. 49, pp. 180-194, Jan. 2001.

[103] Ye Y., "Combining binary search and Newton's method to compute real roots for a class of real functions," *Journal of Complexity*, vol. 10, pp. 271-280, Sep. 1994.

[104] Yoon Y. C. and Leib H., "Maximizing SNR in improper complex noise and application to CDMA," *IEEE Communications Letters*, vol. 1, pp. 5-8, Jan. 1997.

[105] Yuan J.-T. and Lee J.-N., "Narrow-band interference rejection in DS/CDMA systems using adaptive (QRD-LSL)-based nonlinear ACM interpolators," *IEEE Transactions on Vehicular Technology*, vol. 52, pp. 374-379, March 2003.

[106] Yuen N. and Friedlander B., "Asymptotic performance analysis for signature waveform estimation in synchronous CDMA systems," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1753-1757, June 1998.

[107] Zarifi K. and Gershman A. B., "Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems–Part I: General SNR framework," submitted to *IEEE Transactions on Signal Processing*.

[108] Zarifi K. and Gershman A. B., "Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems–Part II: High SNR analysis," submitted to *IEEE Transactions on Signal Processing*.

[109] Zarifi K. and Gershman A. B., "Blind subspace-based signature waveform estimation in BPSK-modulated DS-CDMA systems with circular noise," *IEEE Transactions on Signal Processing*, vol. 54, pp. 3592-3602, Sep. 2006.

[110] Zarifi K. and Gershman A. B., "Blind subspace-based signature estimation in DS-CDMA systems with unknown wide-sense stationary interference," accepted for publication in *IEEE Transactions on Signal Processing*, to appear in July 2007.

[111] Zarifi K. and Gershman A. B., "Enhanced blind subspace-based signature waveform estimation in CDMA systems with circular noise," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, USA, March 2005, pp. 937-940.

[112] Zarifi K. and Gershman A. B., "High SNR performance analysis of blind minimum output energy receivers in large DS-CDMA systems," accepted for *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Honolulu, USA, Apr. 2007.

[113] Zarifi K. and Gershman A. B., "Large system performance analysis of minimum output energy receivers for DS-CDMA systems," submitted to *IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Helsinki, Finland, June 2007.

[114] Zarifi K. and Gershman A. B., "Performance analysis of blind subspace-based signature estimation algorithms for DS-CDMA systems with unknown correlated noise," accepted for publication in *EURASIP Journal on Applied Signal Processing*, special issue on *Advances in Subspace-Based Techniques for Signal Processing and Communications*, to appear in 2007.

[115] Zarifi K. and Gershman A. B., "Performance analysis of subspace-based signature waveform estimation algorithms for DS-CDMA systems with unknown correlated noise," in *Proc. IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, New York, USA, June 2005, pp. 600-604.

[116] Zarifi K. and Gershman A. B., "Subspace-based blind channel estimation in DS-CDMA systems with unknown wide-sense stationary interference," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006, pp. 121-124.

[117] Zarifi K., Shahbazpanahi S., Gershman A. B., and Luo Z.-Q., "Robust blind multiuser detection based on the worst-case performance optimization of the MMSE receiver," *IEEE Transactions on Signal Processing*, vol. 53, pp. 295-305, Jan. 2005.

[118] Zarifi K., Shahbazpanahi S., Gershman A. B., and Luo Z.-Q., "Robust blind multiuser detection based on worst-case MMSE performance," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, May 2004, pp. 897-900.

[119] Zoubir A. M. and Lane-Glover A. T., "Multiuser detection in impulsive noise," in *Proc. IEEE International Workshop on Statistical Signal Processing (SSP)*, Singapore, Aug. 2001, pp. 102-105.

Résumé

Education

- 04.2005–01.2007, Research Associate and Ph.D. Candidate, Department of Communication Systems, Darmstadt University of Technology, Darmstadt, Germany.
- 01.2002–04.2005, Research Associate and Ph.D. Candidate, Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany.
- 09.2002–03.2003, Visiting Researcher, Communication Research Laboratory, McMaster University, Hamilton, Ontario, Canada.
- 09.1997–09.2000, M.Sc. in Electrical Engineering, University of Tehran, Tehran, Iran. *Thesis*: "Robust control of nonlinear systems using nonlinear H_∞ method."
- 09.1993–09.1997, B.Sc. in Electrical Engineering, University of Tehran, Tehran, Iran.

Publications

1. K. Zarifi and A. B. Gershman, "Blind subspace-based signature estimation in DS-CDMA systems with unknown wide-sense stationary interference," accepted for publication in *IEEE Transactions on Signal Processing*, to appear in July 2007.
2. K. Zarifi and A. B. Gershman, "Performance analysis of blind subspace-based signature estimation algorithms for DS-CDMA systems with unknown correlated noise," accepted for publication in *EURASIP Journal on Applied Signal Processing*, special issue on *Advances in Subspace-Based Techniques for Signal Processing and Communications*, to appear in 2007.
3. K. Zarifi and A. B. Gershman, "Blind subspace-based signature waveform estimation in BPSK-modulated DS-CDMA systems with circular noise," *IEEE Transactions on Signal Processing*, vol. 54, pp. 3592-3602, Sep. 2006.
4. K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z.-Q. Luo, "Robust blind multiuser detection based on the worst-case performance optimization of the MMSE receiver," *IEEE Transactions on Signal Processing*, vol. 53, pp. 295-305, Jan. 2005.
5. K. Zarifi and A. B. Gershman, "Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems–Part I: General SNR framework," submitted to *IEEE Transactions on Signal Processing*.

6. K. Zarifi and A. B. Gershman, "Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems–Part II: High SNR analysis," submitted to *IEEE Transactions on Signal Processing*.

7. K. Zarifi and A. B. Gershman, "Large system performance analysis of minimum output energy receivers for DS-CDMA systems," submitted to *IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, June 2007, Helsinki, Finland.

8. K. Zarifi and A. B. Gershman, "High SNR performance analysis of blind minimum output energy receivers in large DS-CDMA systems," accepted for *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2007, Honolulu, USA.

9. K. Zarifi and A. B. Gershman, "Subspace-based blind channel estimation in DS-CDMA systems with unknown wide-sense stationary interference," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2006, Toulouse, France. Finalist in the "**Student Paper Contest**".

10. K. Zarifi and A. B. Gershman, "Performance analysis of subspace-based signature waveform estimation algorithms for DS-CDMA systems with unknown correlated noise," *IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, June 2005, New York, USA.

11. K. Zarifi and A. B. Gershman, "Enhanced blind subspace-based signature waveform estimation in CDMA systems with circular noise," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2005, Philadelphia, USA.

12. K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z.-Q. Luo, "Robust blind multiuser detection based on worst-case MMSE performance," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2004, Montreal, Canada.

13. K. Zarifi and M. J. Yazdanpanah, "Two approximate solutions to HJI equation via state dependent Riccati equation," *European Control Conference*, Sep. 2001, Porto, Portugal.